CUK-H06336-8-261122 C.U.

The British Journal for the Philosophy of Science

CALCUTTA UNIVERSITY CENTRAL LIBRARY

BPS

501.05
B277

Thomas Nelson and Sons Ltd
Edinburgh and London
May 1965

Feb -66

Volume XVI No. 61 Ten Shillings and Sixpence net

IV

MIND

A Quarterly Review of Psychology and Philosophy

edited by Professor **GILBERT RYLE**

with the co-operation of Prof. Sir F. C. Bartlett *and* Prof. C. D. Broad

CONTENTS FOR ISSUE OF JULY 1965

Category Mistakes and Rules of Language	BERNARD HARRISON
Emotion	GEORGE PITCHER
Counterfactuals and Subjunctive Conditionals	M. R. AYERS
Abstract Entities and Universals	MANLEY THOMPSON
Aristotle's Syllogistic and the Fourth Figure	LYNN E. ROSE
Time, Truth and Modalities	KEITH LEHRER and RICHARD TAYLOR
Facts and Rudeness	ANNE LLOYD THOMAS

Discussions:

The Gödel Formula: Some Reservations	RICHARD BUTRICK
The Parenthetical Use of the Verb "Believe"	M. J. CHARLESWORTH
Mr. Eric Toms on the Liar Paradox	SIBAJIBAN
On Quine's "So-called Paradox"	J. M. CHAPMAN and R. J. BUTLER
Mr. Geach on Distribution	ERIC TOMS
The Prediction Paradox Again	JAMES KIEFER and JAMES ELLISON
A Note on Informal Fallacies	RICHARD COLE

NEW BOOKS NOTES

Annual Subscription 25s, post free; single copy 7s 6d, plus postage

Published for MIND ASSOCIATION by

THOMAS NELSON & SONS LTD Parkside Works EDINBURGH 9

JOURNAL OF THE HISTORY OF IDEAS

A Quarterly devoted to Intellectual History

VOLUME XXVI JULY-SEPT., 1965 NUMBER 3

TABLE OF CONTENTS

ARTICLES

Diego Hurtado de Mendoza and Averroism	*Erika Spivakowsky*
Views of Time in Shakespeare	*Ricardo J. Quinones*
Scientific and Historical Revolutions	*G. Wylie Sypher*
Locke's Doctrine of Signs	*Robert L. Armstrong*
Kierkegaard and Leibniz	*Ronald Grimsley*
Hobson Revisited	*Harvey Mitchell*

NOTES

Locke's Epistemology	*Douglas Odegard*
Herder's Russia	*Wolfgang Gesemann*

REVIEW

Review of Jones: *O Strange New World*	*Henry Nash Smith*

Subscription Rates (outside U.S.A. post paid): $6.50 for one year;
$1.75 for single issue

Payable to

JOURNAL OF THE HISTORY OF IDEAS, INC.
The City College, New York 10031, N.Y.

CALCUTTA UNIVERSITY
CENTRAL LIBRARY

The British Journal for the Philosophy of Science

VOLUME XVI No. 61 MAY 1965

CONTENTS

ARTICLES

…CT OF ESSENTIALISM … TAXONOMY – TWO THOUSAND YEARS …IS (II) 1
David L. Hull

GIN AND CONCEPT OF RELATIVITY (II) 19
G. H. Keswani

A MACHINE BE CONSCIOUS? 33
Dennis Thompson

THE DUAL NATURE OF LIGHT 44
George Magyar

DISCUSSIONS

O FORMS OF THE PREDICTION PARADOX 50
B. Meltzer and I. J. Good

LOGIC AND PSYCHOLOGY OF SCIENCE: REPLY TO PROFESSOR KAPP 51
E. H. Hutten

CONTENTS—(*contd.*)

REVIEWS

What is History?
article-review by Burleigh Taylor Wilkins 53

The Anatomy of Inquiry. By Israel Scheffler
reviewed by J. L. Mackie 57

Foresight and Understanding. By Stephen Toulmin
reviewed by John Watling 61

The Development of Logic. By William and Martha Kneale
reviewed by G. T. Kneebone 63

The Principles of Science. A Treatise on Logic and Scientific Method. By W. Stanley Jevons
reviewed by William Kneale

Historical Studies in the Language of Chemistry. By M. P. Crosland
reviewed by John Bradley

Abstracts

Recent Publications on the Philosophy of Science

© Thomas Nelson and Sons Ltd 1965

BRITISH SOCIETY FOR THE PHILOSOPHY OF SCIENCE

The British Society for the Philosophy of Science is a body of which the purpose is described by the following statement from its Constitution and Rules:

> 'The purpose of the Society is to study the logic, the method and the philosophy of science as well as those of the various special sciences, including the social sciences.'

The Society holds meetings, at which papers are read and discussed, once a month during the academic year. These are held, by kind permission, in University College, London.

Ordinary membership is open to all persons who have been proposed by two members, approved by the Committee, and elected at an ordinary meeting. The annual subscription is 10s. 6d. Those who wish to join should write to the Hon. Secretary, Mr. J. W. N. Watkins, London School of Economics, Houghton Street, London, W.C.2.

The Northern Group of the British Society for the Philosophy of Science meets regularly at various centres in the north of England. Those who wish to join it should write to Dr W. Mays, Hon. Secretary of the Northern Group of the British Society for the Philosophy of Science, Dept. of Philosophy, The University, Manchester, 13.

The Present Committee of the Society is as follows:

Professor S. Körner
(President)

Professor W. C. Kneale
(Vice-President)

Mr J. W. N. Watkins
(Hon. Secretary)

Mr G. Buchdahl
(Hon. Treasurer)

Dr C. W. Kilmistre

Dr I. Lakatos

Dr J. S. Wilkie
Mr R. Harré

Dr W. Mays
(representing Northern Group)

THE BRITISH JOURNAL FOR THE PHILOSOPHY OF SCIENCE

The British Journal for the Philosophy of Science is the official organ of the Society. It is published quarterly (February, May, August and November) by Messrs Thomas Nelson & Sons Ltd; it may be obtained from any bookseller or direct from the publisher. The annual subscription is 42s. ($ 6.25), post free; single issues 10s. 6d., postage 6d. The subscription is 31s. 6d. for members of the British Society for the Philosophy of Science, who should order direct from the publishers.

Contributions for inclusion in the *Journal* should no longer be sent to the present Editor, Dr J. O. Wisdom, but to the incoming Editor, Dr Mary Hesse, University of Cambridge, Whipple Museum, Free School Lane, Cambridge. Articles should normally not exceed 5000 words in length. With an article a summary of about 100 words (which should also be used as an abstract) should be sent. **A sheet giving instructions about lay-out, style of references, etc., can be obtained from the Editor on request.**

Exchange of Abstracts. When a paper is accepted for publication, the abstract will then be published in *Dialectica, Methodos,* and *Philosophy of Science.* Likewise, abstracts of papers in these Journals will be found in our pages.

Publishers are invited to forward copies of books to the Editor, *The British Journal for the Philosophy of Science,* c/o Thomas Nelson & Sons Ltd, Parkside Works, Edinburgh 9.

Particulars regarding advertising in the Journal may be obtained from Messrs Thomas Nelson & Sons, 36 Park Street, Park Lane, W.1. (Telephone: HYDe Park 8351.)

The British Journal for the Philosophy of Science

VOLUME XVI No. 62 AUGUST 1965

CONTENTS

ARTICLES

MAN, BEAST, AND PHILOSOPHICAL PSYCHOLOGY 81
by John King-Farlow and Elton A. Hall

THEORETICAL MODELS 102
by Peter Achinstein

MODELS AND THEORIES 121
by Marshall Spector

TOWARDS A THEORETICAL CONCEPTUALISATION OF SUPERSTITION 143
by Karl E. Scheibe and Theodore R. Sarbin

REVIEWS

Method in the Physical Sciences. By G. Schlesinger 159
reviewed by B. C. Brookes

Models and Analogies in Science. By Mary B. Hesse 161
reviewed by Robert Ackermann

CONTENTS—(*contd.*)

REVIEWS

The Minds of Robots. By James T. Culbertson 163
reviewed by D. M. Mackay

Anaxagoras and the Birth of Scientific Method. By D. E. Gershenson and D. A. Greenberg 165
reviewed by E. H. Hutten

ABSTRACTS 167

ANNOUNCEMENT 168

© Thomas Nelson and Sons Ltd 1965

BRITISH SOCIETY FOR THE PHILOSOPHY OF SCIENCE

The British Society for the Philosophy of Science is a body of which the purpose is described by the following statement from its Constitution and Rules:

> 'The purpose of the Society is to study the logic, the method and the philosophy of science as well as those of the various special sciences, including the social sciences.'

The Society holds meetings, at which papers are read and discussed, once a month during the academic year. These are held, by kind permission, in University College, London.

Ordinary membership is open to all persons who have been proposed by two members, approved by the Committee, and elected at an ordinary meeting. The annual subscription is 10s. 6d. Those who wish to join should write to the Hon. Secretary, Mr J. W. N. Watkins, London School of Economics, Houghton Street, London, W.C.2.

The Northern Group of the British Society for the Philosophy of Science meets regularly at various centres in the north of England. Those who wish to join it should write to Dr W. Mays, Hon. Secretary of the Northern Group of the British Society for the Philosophy of Science, Dept. of Philosophy, The University, Manchester, 13.

The Present Committee of the Society is as follows:

Professor S. Körner
(President)

Professor W. C. Kneale
(Vice-President)

Mr J. W. N. Watkins
(Hon. Secretary)

Mr G. Buchdahl
(Hon. Treasurer)

Dr C. W. Kilmister

Dr I. Lakatos

Dr J. S. Wilkie
Mr R. Harré

Dr W. Mays
(representing Northern Group)

THE BRITISH JOURNAL FOR THE PHILOSOPHY OF SCIENCE

The British Journal for the Philosophy of Science is the official organ of the Society. It is published quarterly (February, May, August and November) by Messrs Thomas Nelson & Sons Ltd; it may be obtained from any bookseller or direct from the publisher. The annual subscription is 42s. ($ 6.25), post free; single issues 10s. 6d., postage 6d. The subscription is 31s. 6d. for members of the British Society for the Philosophy of Science, who should order direct from the publishers.

Contributions for inclusion in the *Journal* should be sent to the Editor, Dr Mary Hesse, University of Cambridge, Whipple Museum, Free School Lane, Cambridge. Articles should normally not exceed 5000 words in length. With an article a summary of about 100 words (which should also be used as an abstract) should be sent. **A sheet giving instructions about lay-out, style of references, etc., can be obtained from the Editor on request.**

Exchange of Abstracts. When a paper is accepted for publication, the abstract will then be published in *Dialectica*, *Methodos*, and *Philosophy of Science*. Likewise, abstracts of papers in these Journals will be found in our pages.

Publishers are invited to forward copies of books to the Editor, *The British Journal for the Philosophy of Science*, c/o Thomas Nelson & Sons Ltd, Parkside Works, Edinburgh 9.

Particulars regarding advertising in the Journal may be obtained from Messrs Thomas Nelson & Sons, 36 Park Street, Park Lane, W.1. (Telephone: HYDe Park 8351.)

The British Journal for the Philosophy of Science

VOLUME XVI No. 63 NOVEMBER 1965

CONTENTS

ARTICLES

THE LOGIC OF RATIONAL DECISION 169
by Herbert A. Simon

CAUSALITY, CAUSAL LAWS AND SCIENTIFIC THEORY IN THE PHILOSOPHY OF KANT 187
by Gerd Buchdahl

CONNECTIVITY, CHANCE, AND IGNORANCE 209
by D. H. Mellor

THE CONCEPT OF PROBABILITY 226
by Neil Cooper

DISCUSSIONS

IN DEFENCE OF THE SOCIOLOGY OF KNOWLEDGE 239
by Irving Louis Horowitz

REPLY TO PROFESSOR HOROWITZ 241
by I. C. Jarvie

CONTENTS—(*contd.*)

NOTE ON MR KESWANI'S ARTICLES, ORIGIN AND CONCEPT OF RELATIVITY 242
by Herbert Dingle

THE ORIGIN AND CONCEPT OF RELATIVITY 246
by H. B. Levinson

MAN, BEAST, AND PHILOSOPHICAL PSYCHOLOGY 246
by Peter Geach

REVIEWS

Mandelbaum's Critical Realism 249
article-review by J. W. N. Watkins

Language and Information—Selected Essays on Their Theory and Application. By Y. Bar-Hillel 253
reviewed by D. M. MacKay

The Act of Creation. By Arthur Koestler 255
reviewed by Richard Robinson

ABSTRACTS 258

© Thomas Nelson and Sons Ltd 1965

BRITISH SOCIETY FOR THE PHILOSOPHY OF SCIENCE

The British Society for the Philosophy of Science is a body of which the purpose is described by the following statement from its Constitution and Rules:

> 'The purpose of the Society is to study the logic, the method and the philosophy of science as well as those of the various special sciences, including the social sciences.'

The Society holds meetings, at which papers are read and discussed, once a month during the academic year. These are held, by kind permission, in University College, London.

Ordinary membership is open to all persons who have been proposed by two members, approved by the Committee, and elected at an ordinary meeting. The annual subscription is 10s. 6d. Those who wish to join should write to the Hon. Secretary, Mr [illegible]. W. N. Watkins, London School of Economics, Houghton Street, London, W.C.2.

The Northern Group of the British Society for the Philosophy of Science meets regularly at various centres in the north [illegible] England. Those who wish to join it should write to Dr W. Mays, Hon. Secretary of the Northern Group of the British Society for the Philosophy of Science, Dept. of Philosophy, The University, Manchester, 13.

The Present Committee of the Society is as follows:

Professor S. Körner (*President*)	Dr C. W. Kilmister
Professor W. C. Kneale (*Vice-President*)	Dr I. Lakatos
Mr J. W. N. Watkins (*Hon. Secretary*)	Dr J. S. Wilkie Mr R. Harré
Mr G. Buchdahl (*Hon. Treasurer*)	Dr W. Mays (*representing Northern Group*)

THE BRITISH JOURNAL FOR THE PHILOSOPHY OF SCIENCE

The British Journal for the Philosophy of Science is the official organ of the Society. It is published quarterly (February, May, August and November) by Messrs Thomas Nelson & Sons Ltd; it may be obtained from any bookseller or direct from the publisher. The annual subscription is 42s. ($ 6.25), post free; single issues 10s. 6d., postage 6d. The subscription is 31s. 6d. for members of the British Society for the Philosophy of Science, who should order direct from the publishers.

Contributions for inclusion in the *Journal* should be sent to the Editor, Dr Mary Hesse, University of Cambridge, Whipple Museum, Free School Lane, Cambridge. Articles should normally not exceed 5000 words in length. With an article a summary of about 100 words (which should also be used as an abstract) should be sent. **A sheet giving instructions about lay-out, style of references, etc., can be obtained from the Editor on request.**

Exchange of Abstracts. W[illegible] paper is accepted for the publication, the abstract will then be publ[illegible] in *Dialectica*, *Methodos*, and *Philosophy of Science*. Likewise, abstracts of papers in these Journals will be found in our pages.

Publishers are invited to forward copies of books to the Editor, *The British Journal for the Philosophy of Science*, c/o Thomas Nelson & Sons Ltd, Parkside Works, Edinburgh 9.

Particulars regarding advertising in the Journal may be obtained from Messrs Thomas Nelson & Sons, 36 Park Street, Park Lane, W.1. (Telephone: HYDe Park 8351.)

CALCUTTA UNIVERSITY
CENTRAL LIBRARY

The British Journal for the Philosophy of Science

VOLUME XVI No 64 FEBRUARY 1966

CONTENTS

ARTICLES

ON LOGICAL DEFINITIONS OF CONFIRMATION 265
by D. Stove

ORIGIN AND CONCEPT OF RELATIVITY (III) 273
by G. H. Keswani

PHYSICS AND COMMON SENSE 295
by Nicholas Maxwell

CONFIRMATORY MODELS OF THEORIES 312
by Robert J. Ackermann

DISCUSSIONS

ESTIMATED UTILITY AND CORROBORATION 327
by Alex. C. Michalos

A NOTE ON THE DIFFERENCE BETWEEN THE LORENTZ-FITZGERALD CONTRACTION AND THE EINSTEIN CONTRACTION 332
by Karl R. Popper

CONTENTS—(*contd.*)

REVIEWS

SUBJECTIVE PROBABILITY 334
article-review by Ian Hacking

Act and Agent: An Essay in Philosophical Anthropology. By Douglas Browning 340
reviewed by J. H. M. Beattie

Signification and Significance: a study of the relation of signs and values. By Charles Morris 341
reviewed by J. D. North

Logic and Language. Edited by B. H. Kazemier and D. Vuysje 344
reviewed by James Cargile

Science, Technology, and Human Values. By A. Cornelius Benjamin 346
reveiwed by C. W. Keenan

My View of the World. By E. Schrödinger 348
reviewed by G. P. Thomson

Contributions to Logic and Methodology in Honor of J. M. Bochenski. Edited by Anna-Teresa Tymieniecka 349
reviewed by Desmond Paul Henry

Le Concept d'Information dans la Science Contemporaine. Edited by M. Louis Couffignal 351
reviewed by D. Gabor

© Thomas Nelson and Sons Ltd., 1966

BRITISH SOCIETY FOR THE PHILOSOPHY OF SCIENCE

The British Society for the Philosophy of Science is a body of which the purpose is described by the following statement from its Constitution and Rules:

> 'The purpose of the Society is to study the logic, the method and the philosophy of science as well as those of the various special sciences, including the social sciences.'

The Society holds meetings, at which papers are read and discussed, once a month during the academic year. These are held, by kind permission, in University College, London.

Ordinary membership is open to all persons who have been proposed by two members, approved by the Committee, and elected at an ordinary meeting. The annual subscription is 10s. 6d. Those who wish to join should write to the Hon. Secretary, Mr J. W. N. Watkins, London School of Economics, Houghton Street, London, W.C.2.

The Northern Group of the British Society for the Philosophy of Science meets regularly at various centres in the north of England. Those who wish to join it should write to Dr W. Mays, Hon. Secretary of the Northern Group of the British Society for the Philosophy of Science, Dept. of Philosophy, The University, Manchester, 13.

The Present Committee of the Society is as follows:

Professor S. Körner
(President)

Professor W. C. Kneale
(Vice-President)

Mr J. W. N. Watkins
(Hon. Secretary)

Mr G. Buchdahl
(Hon. Treasurer)

Dr C. W. Kilmister

Dr I. Lakatos

Dr J. S. Wilkie
Mr H. D. Sluga

Dr W. Mays
(representing Northern Group)

THE BRITISH JOURNAL FOR THE PHILOSOPHY OF SCIENCE

The British Journal for the Philosophy of Science is the official organ of the Society. It is published quarterly (February, May, August and November) by Messrs Thomas Nelson & Sons Ltd; it may be obtained from any bookseller or direct from the publisher. The annual subscription is 42s. ($ 6.25), post free; single issues 10s. 6d., postage 6d. The subscription is 31s. 6d. for members of the British Society for the Philosophy of Science, who should order direct from the publishers.

Contributions for inclusion in the *Journal* should be sent to the Editor, Dr Mary Hesse, University of Cambridge, Whipple Museum, Free School Lane, Cambridge. Articles should normally not exceed 5000 words in length. With an article a summary of about 100 words (which should also be used as an abstract) should be sent. **A sheet giving instructions about lay-out, style of references, etc., can be obtained from the Editor on request.**

Exchange of Abstracts. When a paper is accepted for publication, the abstract will then be published in *Dialectica*, *Methodos*, and *Philosophy of Science*. Likewise, abstracts of papers in these Journals will be found in our pages.

Publishers are invited to forward copies of books to the Editor, *The British Journal for the Philosophy of Science*, c/o Thomas Nelson & Sons Ltd, Parkside Works Edinburgh 9.

Particulars regarding advertising in the Journal may be obtained from Messrs Thomas Nelson & Sons, 36 Park Street, Park Lane, W.1. (Telephone: HYDe Park 8351.)

The British Journal for the Philosophy of Science

VOLUME XVI MAY, 1965 No. 61

THE EFFECT OF ESSENTIALISM ON TAXONOMY— TWO THOUSAND YEARS OF STASIS (II)*

DAVID L. HULL

5 *The Species Problem*

THUS far all that has been shown is that with respect to taxa names taxonomists have rejected Aristotelian definition. It has yet to be shown that they have failed to eliminate completely their predisposition for Aristotelian definition and that this failure has been at least in part responsible for the persistence of the species problem. Species will be treated in this paper only from the point of view of phylogenetic taxonomy; that is, from the point of view that classification must have some systematic relationship to phylogeny and that the unit of classification must be the unit of evolution. In order for such a position on the status of 'species' to be justified, not only must an adequate definition of 'species' as an evolutionary unit be given but also the phylogenetic programme itself must be shown to be feasible and significant. Only the former will be attempted in this paper.

G. G. Simpson defines an evolutionary species as 'a lineage (an ancestral-descendant sequence of populations) evolving separately from others and with its own unitary evolutionary role and tendencies'.[1] Unfortunately, Simpson's definition itself does not provide any explicitly formulated criteria for determining exactly how unitary an evolutionary role is unitary enough for species status. He defines roles in terms of niches and niches in terms of whole ways of life. Simpson's definition is important because it brings to the fore one of the basic principles of phylogenetic taxonomy (that the unit of classification must be the unit of evolution), but it does not contribute much toward determining what level of evolutionary unity is to be considered specific. In short it is not ' operational '. The purpose of the following sections will be to provide criteria to implement Simpson's definition.

---

* Part I appeared in the previous number.

[1] Simpson, 1961, p. 153

Dobzhansky provided one such criterion in his biological definition. When a group of organisms which usually reproduce by interbreeding interbreeds consistently and produces reasonably fertile offspring, this group of organisms is as distinct an evolutionary unit as there is in phylogeny. The members of the group are *genetically* affecting each other's phylogenetic development. But the biological definition is not without its faults. Cain summarises the major difficulties confronting the criterion of actual interbreeding as follows: ' The biospecies is a definable concept only if time and allopatric populations are ignored and asexual forms are excluded from consideration '.[1] These are not the *only* difficulties confronting a successful definition of ' species ', but they are certainly three of the major ones. If they can be surmounted, the others will be relatively easy to account for.

Before each of these major difficulties is treated, reference must be made to Dobzhansky's second criterion for species status. Dobzhansky was well aware that actual interbreeding applied to only a small percentage of organisms. (Our ability to observe such actual interbreeding directly applies to an even smaller percentage.) To accommodate all other cases (that is, to accommodate a majority of the cases) he introduced the criterion of potential interbreeding. In doing so he adopted a non-traditional form of definition, a disjunctive definition. He did not, however, depart very radically from traditional definition, since the two disjuncts were considered to be the only pertinent conditions. Each was sufficient, the fulfilment of at least one necessary. Dobzhansky's definition has much to recommend it. One of the criteria is at least the basis for a sufficient condition for delineating species as units of evolution. The form of the definition, although not completely adequate, is more appropriate than the form of all previous definitions. Unfortunately, Dobzhansky chose potential interbreeding as his second criterion.

If it had not other faults, the criterion of potential interbreeding would be undesirable on the grounds of vagueness alone. It is intended to cover too many too varied exceptions to the first criterion.

---

[1] Cain, 1954, p. 24

In one sense a population of mice on one island is potentially interbreeding with a population of mice on another island. They would interbreed if they could get at each other. In a second sense a population of Drosophila living in 1942 is potentially interbreeding with a population of that species living in 1922 or 1962. If they had lived at the same time in the same area, they would have interbred. In still another sense, two interfertile sibling species which are both sympatric and synchronic are potentially interbreeding. They could interbreed if they were so inclined. In another sense all the breeds of domestic dogs are potentially interbreeding. They would (and do) interbreed whenever their masters permit. The list could be extended indefinitely.

There is, however, a more fundamental philosophical reason for abandoning potential interbreeding as a criterion for species status. Bertrand Russell says, for example:

> The concept of potentiality is convenient in some connections, provided it is so used that we can translate our statements into a form in which the concept is absent. 'A block of marble is a potential statue' means 'from a block of marble, by suitable acts, a statue is produced'. But when potentiality is used as a fundamental and irreducible concept it always conceals confusion of thought. Aristotle's use of it is one of the bad points in his system.[1]

Whether or not such a blanket indictment of the concept of potentiality is completely justified, there are reasons peculiar to evolutionary taxonomy for avoiding it. What is important in evolution is not which organisms could interbreed but which organisms do interbreed. The fact that two groups of organisms *cannot* interbreed (regardless of the isolating mechanism) is important only in the respect that it follows deductively that they *are not* interbreeding. On the other hand the fact that two groups of organisms can or could interbreed even though they are not so interbreeding (regardless of how this is to be ascertained) is important in only two respects. First, it means that the two groups of organisms have not diverged appreciably from each other in interbreeding habits since they had a common ancestry which was actually interbreeding. Second, it means that if in the future the isolating mechanisms are removed, then the two groups would interbreed. But taxonomists are not obliged to predict the future course of evolution. Taxonomists are obliged to classify only those species that have evolved given the environment that did pertain, not to classify all possible species that might have evolved in some possible environment. Until potentially interbreeding organisms actually use this potentiality, it is of only ' potential ' interest in classification. In evolutionary taxonomy *unrealised* potentialities don't count.

---

[1] Bertrand Russell, *A History of Western Philosophy*, New York, 1945, p. 167

A. *Allopatric Populations*

Synchronic populations can be geographically separated in two ways. Either they can be separated but connected by intermediate populations or else they can be completely isolated from each other. Chains of contiguous or overlapping populations are termed geographic *Rassenkreise*. Such chains may vary with respect to any property or type of property, but the two most important kinds of *Rassenkreise* are those that vary with respect to morphological similarity and those that vary with respect to interbreeding. Douglas Gasking defines a morphological species as all those forms that are serially very like each other[1] and a biological species as all those populations that are serially crossable with each other.[2] A serial relation is a relation such that the simple relation holds between any two consecutive terms in the series but need not hold between any two terms that are not consecutive.

For example in a morphological *Rassenkreis*, population *A* may be very like population *B*, which is very like population *C*, which is very like population *D*, and so on, but *A* and a distant population, say, *G* may not be very like each other at all. Such continuous variation among contemporary populations presents no problem for the biological definition because interbreeding status is the criterion for species status among contemporary populations, not morphological similarity. However, interbreeding *Rassenkreise* do present a minor difficulty. In some cases population *A* is interbreeding with population *B*, *B* with *C*, *C* with *D* and so on; but not only is it the case that *A* is not interbreeding with a distant member of the chain, but also it is sometimes the case that they are not even interfertile. Nevertheless, genes are exchanged via intermediate populations. All members of the interbreeding *Rassenkreis* belong to the same ' gene pool ' and are evolving with a sufficient degree of separateness and unity to be classed as species. They are genetically affecting each other's evolutionary development.

The second respect in which two synchronic populations can be separated geographically is by complete isolation. In such cases the populations are neither interbreeding nor serially interbreeding, and yet taxonomists class these isolated populations in the same species if they are morphologically similar to each other and distinct from other species; that is, if they possess a degree of morphological similarity and difference usually indicative of species status in organisms of that type. The degree of morphological similarity and difference usually indicative of species status is determined by the degree of morphological similarity and difference present among contemporary interbreeding species.

---

[1] Douglas Gasking, ' Clusters ', *Australasian Review of Psychology*, 1960, 38, 13, 18

[2] Gasking, 1960, p. 38

The question of greatest interest to the biologist is how good morphological similarity and difference is at indicating species status if species are to be units of evolution. Although this is primarily an empirical question and can be answered only after extensive empirical investigation, its solution has been hindered by logical and philosophical confusions. The phylogeneticists themselves have been responsible for some of the confusion by treating consistent interbreeding with the production of fertile offspring as both a necessary and a sufficient condition when it is obviously only sufficient. For example, true sibling species have been extremely troublesome to the phylogeneticists. From a practical standpoint, taxonomists would like to treat them as single species although they do not interbreed even when given every opportunity. On the other hand, from a theoretical point of view, classing them as a single species seems to run counter to the biological definition.[1] Classing two groups of organisms which are morphologically and ecologically quite similar as a single species although they never interbreed conflicts with the biological definition only if interbreeding is mistakenly considered to be a necessary condition. As Simpson has pointed out, a pair of sibling species which are distinguishable neither morphologically nor ecologically are evolving as an evolutionary unit—albeit without the aid of genetic interchange.[2] As will be seen shortly, the same can be said for species of asexual organisms.

A second confusion must be laid at the doorstep of the opponents of phylogenetic taxonomy. Although the point has been made often and well, it bears repeating: it does not follow from the fact that morphological similarity and difference are used as the evidence from which species status is inferred that the morphological definition of 'species' has been substituted for the biological definition. Morphological similarity and difference is only the *evidence* being used to determine species status. How similar is similar enough is determined by the criterion of interbreeding and *this* is what is logically important.

---

[1] Mayr (1957), p. 376; note Mayr, 1942, p. 200, for his treating interbreeding as if it were a necessary condition at least with respect to sibling species.

[2] Simpson, 1961, p. 160

Finally, A. J. Cain voices a common complaint against inferring species status from morphological similarity when he says, 'We can determine by observation the specific or merely varietal status of partially or completely sympatric forms, but can only guess at the status of wholly allopatric ones'.[1] Several factors contribute to this view, including such basic philosophical problems as the justification of induction and the role of deduction in science. The inferences taxonomists make from morphological similarity to species status are not readily put into a form which lends itself to easy manipulation within any of the various theories of probability which have presently been devised. This fact says more against the present development of probability theory than it does against such scientific inferences. Most of the inferences scientists make cannot be accounted for in probability theory. It might also be fruitfully mentioned that the most scathing critics of inferring that two populations belong to the same species from morphological similarity advocate a comparable inference that two different instances of a property are instances of the same property.[2]

The importance of justifying the inference from morphological similarity to phylogenetic relationship should not be underestimated. If it is not justified in a good percentage of the cases, then the entire phylogenetic programme becomes untenable since in most cases morphological properties are *all* the taxonomist has to go on. But before such a harsh judgment is passed, one should make sure that phylogenetic taxonomy is not being measured against an unrealistic standard. On the one hand, science is not as empirical as many scientists seem to think it is. Unobserved and even unobservable entities play an important part in it. Science is not just the making of observations: it is the making of inferences on the basis of observations within the framework of a theory. On the other hand, most of the inferences made by scientists are not deductions as many logicians and philosophers seem to think they are. All inferences made by scientists need not match the accuracy possible in certain restricted areas of physics to be justified.[3] Inductive inferences are not deductively certain.

---

[1] Cain, 1954, p. 73

[2] For example, Sokal and Sneath, 1963. Either both types of inference are justified or else neither is.

[3] Michael J. Scriven, 'The Key Property of Physical Laws—Inaccuracy', H. Feigl and G. Maxwell (eds.), *Current Issues in the Philosophy of Science*, New York, 1961

If the inferences made by phylogeneticists are guessing, then so are those of meteorologists, economists, historians, pollsters and so on. According to the definition of 'guessing' implied by Cain's comment, most of what is known as science, including all of the social sciences, becomes guessing. Perhaps the scepticism with which Cain and others view inference from morphological similarity to species status arises from the confusion wrought by the criterion of potential interbreeding. If the status which taxonomists are to determine for wholly allopatric forms is whether or not they are potentially interbreeding, then perhaps the critics are justified in terming such inference 'guessing'. If taxonomists are expected to predict the future development of the organisms being classified, then they are guessing in the strictest sense of the word, since both gene and evolutionary theory are not predictive but retrodictive theories. If on the other hand taxonomists are expected only to infer what has actually happened, what species have actually evolved, what groups were actually interbreeding, then these inferences are well outside the range of guessing.

In any case the purpose of this paper is to present a *type* of definition appropriate for evolutionary taxonomy. Even if the opponents of phylogenetic taxonomy can show that phylogenetic relationships cannot be inferred with reasonable accuracy from the type and extent of evidence that the phylogeneticists have at their disposal, this fact will have no bearing on the *logical* assertion which is the thesis of this paper that Aristotelian definition in terms of a set of necessary and sufficient conditions is inadequate for defining 'species' if species are to be the units of evolution.

B. *Allochronic Populations*

Taxonomists are unanimous in their opinion that temporal isolation presents a more serious problem for the biological definition of 'species' than geographic isolation presents. A. J. Cain says, for example:

> The palaeospecies is an expression of the attempted imposition of a hierarchy developed for classifying discrete groups, on to a continuous evolutionary series. Because of the imperfections of the fossil record many fossils do fall into morphologically discrete groups and can readily be incorporated into the hierarchy. Nevertheless, the whole concept of the species as a morphologically (and by implication genetically) discrete group is based upon the observation of present-day animals, and holds only for short periods of time which on the evolutionary scale are mere instants.[1]

---

[1] Cain, 1954, p. 123

Evolving lineages form what might be called temporal *Rassenkreise* both with respect to morphological similarity and with respect to interbreeding. Each successive generation in a progressively evolving lineage changes only slightly. Each generation is serially very like all of its ancestral generations. Similarly, each generation of interbreeding forms is serially interbreeding with all of its ancestral generations. Thus, ' species ' cannot be defined in terms of a serial relation unless a temporal restriction is imposed on pain of classifying all organisms in a single species.

Morphological *Rassenkreise* among contemporary forms presented no problems for the biological definition because interbreeding status could be determined directly. In temporally separated populations, however, whether there is or is not a complete fossil record, interbreeding status cannot be determined directly. Yet phylogeneticists must find some way to divide progressively evolving lineages into evolutionary units. In the case of lineages broken by fossil gaps, the task is easy. In the case of lineages for which there is a reasonably complete fossil record, the task is not so easy. Even so, there is a solution to the problem of dividing progressively evolving lineages objectively and non-arbitrarily, and the key to the solution is again interbreeding.

The importance of interbreeding for determining species status has been emphasised time and again, but the extent of the criterion's significance has not been fully appreciated even by its strongest proponents. Mayr, using a suggestive metaphor, compares the biological definition to a ' yardstick ' for determining species status, and he rightly maintains that division by such a yardstick is both objective and non-arbitrary.[1] Further, he sees that what difficulties there are concerning the biological definition are not ' with the yardstick but with its application '.[2] A further development of Mayr's spatial analogy comparable to Simpson's analogy of a string gradually changing colour proves quite enlightening in the determination of the roles of interbreeding and morphological similarity and difference and placing them in their proper perspective.

The paradigm case of objective measurement is the measurement of objects in space or of the distances between spatially separated objects. Space as such is an amorphous continuum with no intrinsic metric.[3] What this means is that there is nothing about space itself to indicate

[1] Mayr, 1957, pp. 15, 16, 18, 19

[2] Mayr, 1959, p. 375

[3] Adolf Grünbaum, *Philosophical Problems of Space and Time*, New York, 1963

how long to make the unit of measurement or at what point to begin measuring once the unit of length has been chosen. With respect to objects in space, the choice of the unit of length is primarily a matter of convenience. Bacteria are measured in microns, rugs in feet (or yards or metres), roads in miles or kilometres, and interstellar distances in light years. One highly advantageous property for a system of measurement is to have the units commensurable. A mile can be divided evenly into yards, feet, or inches. At what point we begin to measure is a matter of convenience. We could begin to measure a rug by laying down our ruler anywhere on it, although it is most convenient to begin at one corner and measure one edge. Nothing about space itself dictates either the length of the unit of measurement or the point at which to begin measuring, but it does dictate what type of unit is appropriate—a spatial unit.

Comparable points can be made about what might be called ' taxonomic space ', but it must be emphasised that everything that is said in terms of taxonomic space is meant merely to be explanatory and suggestive. Talk of taxonomic space, like talk of gene pools, is strictly metaphorical. Before either metaphor can actually function in taxonomic theory, a rigorous development in exclusively taxonomic terms is required. Taxonomic space is constructed by plotting ' morphological distance ' on the horizontal axis and development in time on the other. ' Morphological distance ' means morphological similarity and difference. Taxonomic space, like physical space, is an amorphous continuum with no intrinsic metric; there is nothing about taxonomic space to indicate how long to make the unit of measurement or at what point to begin measuring once the unit [illegible] chosen.

With respect to the objects in taxonomic space, the choice of [illegible] unit of length is neither arbitrary nor a matter of choice for the phylogeneticists. (This is not the case for the classificationists and their numerical progeny.) The unit of taxonomic space is the morphological distance usually indicative of interbreeding status among contemporary organisms which usually reproduce by interbreeding. However, interbreeding is not the unit of taxonomic space; interbreeding merely determines the length of the unit of taxonomic space. In terms of Mayr's metaphor, the yardstick is morphological distance. Interbreeding determines how long a yard of morphological space is. Nor is the choice of the point at which to begin measuring lineages arbitrary or purely a matter of convenience. In the case of lineages which terminate in contemporary species, it is certainly most convenient to begin laying down our yardstick with them; but there are also theoretical reasons to begin with contemporary species, for only with them can we check the accuracy of our yardstick. The purpose of the yardstick is to delineate evolutionary units. The rationale for making the yardstick one length rather than another is that for the group of organisms in question a particular morphological distance is usually indicative of interbreeding status, which is indicative that the group is evolving as a unit and is, hence, rightly called a species. The nature of taxonomic space as we constructed it determines what type of unit is appropriate for measuring lineages—morphological distance. In the case of lineages which are entirely extinct, measurement begins at the point of extinction or with the most recent fossil. These points may or may not coincide. In the case of lineages which are known only by isolated groups of fossils, the length of the object to be measured is usually less than the length of the yardstick so no division is necessary.

One rather disconcerting property of both physical and taxonomic space, which has been referred to only tangentially, is that fact that the length of the unit of measurement varies. Just as spatially extended objects determine the contours of physical space, evolutionary lineages determine the topography of taxonomic space. To stretch the analogy to its breaking point, taxonomic space is 'non-euclidean'. The vertical temporal axis is divided evenly into segments, but the horizontal axis which indicates morphological distance is not. The units of morphological distance vary in length depending on what morpho[illegible]gical distance is indicative of species status for the contemporary [illegible]mbers of the general type of organism. Some organisms interbreed [illegible]n if there is great morphological dissimilarity; others permit almost none. If the properties under consideration are weighted, still another dimension of complexity is added. The logic of the situation remains unchanged, however.

Taxonomists use the yardstick of morphological distance to indicate species status among contemporary allopatric populations. There is no reason why its use cannot be extended to allochronic or even to allochronic and allopatric forms.[1] Although Simpson feels that the resulting divisions will be arbitrary, he condones chopping a progressively evolving lineage into segments by use of morphological distance. He goes on to say, however, 'In practice all that is needed is some criterion as to how large (and in what sense "large") to make the

---

[1] Mayr, 1957, p. 120, and Cain, 1954, p. 111

segments.... Successive species should be so defined as to make the morphological differences between them at least as great as sequential differences among contemporaneous species of the same or closely allied groups'.[1] In progressively evolving lineages there is no sudden discontinuity which can be used as a specific boundary, just as there is none in a synchronic morphological *Rassenkreis*; but there *is* a reason in both cases to make a break in the continuous series at one point rather than at another.

C. *Asexual Populations*[2]

Taxonomists universally agree (if universal agreement among taxonomists is possible) that organisms which never reproduce by interbreeding, whether contemporary or ancestral, present an insurmountable difficulty for the biological definition which no amount of modification could accommodate. The following quotations are typical:

> In summary, the existence of the species, as an objective biological unit is not impaired by morphological indistinctness or by the continuity of the evolutionary process. The loss of sexuality, on the other hand, removes the very foundations on which the species exists as a type of breeding population. As a result biological species do not exist in asexual groups.[3]

> The essence of the biological species concept is discontinuity due to reproductive isolation. Without sexuality this concept cannot be applied. Asexuality then is the most formidable and most fundamental obstacle of a biological species concept.[4]

> The agamospecies represents an advance on the morphospecies, since the mode of reproduction is known. Unfortunately, it is a mode which allows no possibility of framing a definition of the species which is any less arbitrary than the morphospecies.[5]

In the case of asexual forms, taxonomists feel they have come to an impasse in applying the yardstick they have devised to determine species status, for it is logically impossible for a non-interbreeding species to be interbreeding (not to *become* interbreeding but to *be* interbreeding). What they have neglected is that the yardstick is not interbreeding status but the degree of morphological distance usually indicative of interbreeding status. When a taxonomist says that a certain asexual species possesses a degree of morphological similarity and difference comparable to that of contemporary interbreeding species, he is not saying something that is logically impossible, although it may be empirically false. Mayr says, 'It is possible to use the same kind of inference to classify asexual organisms into species. Those asexual individuals are included in a single species that display no more morphological difference from each other than do conspecific individuals in related sexual species'.[1]

---

[1] Simpson, 1961, p. 165

[2] Mayr disapproves of calling groups of asexual organisms 'populations'. Simpson has no reservations concerning the term. For example, see his evolutionary definition of 'species' as quoted in this paper.

[3] Verne Grant, 'The plant species in theory and practice'; *The Species Problem*, Washington, 1957, p. 61

[4] Mayr, 1957, p. 379

[5] Cain, 1954, p. 123

Although T. M. Sonneborn wants to retain the word 'species' to refer to the unit of identification and coins the word 'syngen' to refer to the unit of evolution as determined by the biological definition, he says, 'The question of whether asexual equivalents of syngens exist and can be recognized is more difficult, but not hopeless. The key to progress in this direction is to recognize in the syngen of sexual organisms a distinction between the means of ascertainment and that which is ascertained'.[2] He goes on to say in more detail:

> The preceding attempt to generalize the biological species or syngen runs counter to the view of proponents of the biological species concept . . . that biological species do not exist among obligatory inbreeders or asexual organisms. This denial, as indicated above, is based on an operational definition of biological species. Since the operation, testing gene flow, is impossible in asexual organisms, they deny the existence in them of the thing this operation discovers in sexual organisms, i.e., the biological species or syngen. Their statement of the situation thus implies an abrupt change in the organization of nature and in the units of evolutionary divergence correlated with an abrupt change from outbreeding to obligatory inbreeding and asexual reproduction. By subordinating the method of ascertainment to the thing ascertained and by seeking methods of ascertainment in asexual reproduction, the concept of biological species or syngens was generalized.

---

[1] Mayr, 1957, p. 381. Happily organisms which reproduce sexually are scattered throughout the animal kingdom from Protozoa to Chordata enabling taxonomists to establish the morphological yardstick without making unreasonable inductive leaps.

[2] T. M. Sonneborn, 'Breeding Systems, reproductive methods, and species problems in Protozoa', *The Species Problem*, Washington, 1957, p. 290

This implies the absence of an abrupt change in the organization of nature and in the units of evolutionary divergence with changes in breeding systems or method of reproduction.

No such abrupt change is in fact found in the present review of conditions in Protozoa.[1]

E. A Disjunctive Definition of 'Species'

If the yardstick used to determine species status among allochronic and allopatric interbreeding species is accurate enough to indicate evolutionary units, and if it can be extended to apply to asexual species without any great loss in accuracy, then it would seem that the phylogeneticists at least are in a position to formulate a definition of 'species' adequate for their purposes. 'Species' could be defined *disjunctively* as populations that:

(1) consistently interbreed producing a reasonably large proportion of reasonably fertile offspring, or

(2) consistently serially interbreed with synchronic populations producing a reasonably large proportion of reasonably fertile offspring, or

(3) do not fulfil either of the first two conditions but have not diverged appreciably from a common ancestry which did fulfil one of them, or

(4) do not fulfil any of the first three conditions because they do not apply but are analogous to populations which do fulfil at least one of the first three conditions.

Conditions (1) and (2) are quite straightforward and require no application of the yardstick of morphological distance. In fact they are the basis for determining the length of the morphological yardstick. However, they apply only to contemporary sympatric populations of organisms which usually reproduce by interbreeding. Condition (3) applies only to populations which usually reproduce by interbreeding, but there are no temporal or spatial restrictions. How much divergence is 'appreciable' is determined by the morphological yardstick. Condition (4) applies only to non-interbreeding populations with no temporal or spatial restrictions. How analogous the populations have to be is determined again by the morphological yardstick.

As in the case of all disjunctive definitions, the fulfilling of *any one* of the conditions is sufficient and the fulfilling of *at least one* is necessary. Like the disjunctive definitions of taxa names, it is not insulated against empirical considerations, both because the objects in taxonomic space determine the length of the morphological yardstick and because new methods of reproduction might arise or be discovered. For example, some asexual forms exhibit a phenomenon called ' parasexuality ', in which all or some of the genetic material of one individual is transferred to another. If it is found that these non-meiotic mechanisms of gene flow are sufficient to ensure evolutionary unity and separateness, then either an additional condition would have to be added to the disjunctive definition of ' species ' or else the meaning of ' interbreeding ' would have to be expanded to encompass parasexuality. The former is the more probable move.

---

[1] T. M. Sonneborn, 1957, p. 296.

However, instead of taxonomists concluding that ' species ' can be defined, albeit disjunctively, they have concluded just the opposite. ' Species ' cannot be defined because there are several *kinds* of species. Mayr says, for example, ' Two facts emerge from these and other classifications. One is that there is more than one species concept and that it is futile to search for *the* species concept '.[1] The reasons which taxonomists give for refusing to accept a disjunctive definition of ' species ' and the moves they make to avoid defining it disjunctively are exactly the reasons that the traditionalists give and the moves they make for avoiding disjunctive definition of *any* term. According to the advocates of traditional Aristotelian definition, whenever a word can be defined only disjunctively, either one or more conditions must be accepted as necessary and sufficient or else the fuzzy concept must be divided into terms that can be defined traditionally. The first alternative is not open to phylogeneticists given evolutionary development as it is. Sonneborn outlines the possibilities given the second alternative.

> (i) Apply the word ' species ' to all the various kinds of species, including species as the unit of identification and as the unit of evolution.
> (ii) Retain the term to apply to the unit of evolution as determined by the biological definition.
> (iii) Continue to use the term to apply to the unit of identification as determined by the morphological definition.[2]

Sonneborn rejects the first alternative ' on the grounds that a technical term should have a single meaning, and the second alternative on the grounds of priority and generality '.[3] Thus, ' species ' is to be reserved for the units of identification and groups of organisms which fulfil the requirements of the biological definition are to be termed ‘ syngens ’. V. Grant opts for the second alternative and suggests the term ‘ binom ’ for groups of organisms which do not fulfil the requirements of the biological definition.[1] Cain suggests adopting either (i) or (ii). He says, ‘ One can either restrict the name to one meaning (presumably the biological species) or use it with appropriate prefixes for all ’.[2] Regardless of how it may sound, these biologists are *not* arguing about words. ‘ Species ’ has no magic power, and they are well aware of it. What they are arguing about is whether the unit of identification or the unit of evolution will be the unit of classification and, thus, will remain intact as higher taxa are constructed. In order for the purposes of phylogenetic taxonomy to be fulfilled, the unit of evolution must remain intact.[3]

---

[1] Mayr, 1957, p. 10

[2] Sonneborn, 1957, p. 201

[3] Ibid., p. 201

As is usually the case with disjunctive definitions, taxonomists feel two opposing pulls. On the one hand, since several alternative conditions are each sufficient, they feel that several different terms are being defined. Thus, they conclude that there are several different kinds of species. On the other hand, these different kinds of species seem to function with amazing similarity in both evolutionary and gene theory. Almost everything that a biologist would want to say about one kind of species, he would want to say about the other, with the exception of statements directly pertaining to their defining criteria. Thus, G. G. Simpson says, ‘ The evolution of uniparental and biparental populations is different in many important ways. That does not alter the fact that both form species and, by appropriate definition, the same kind of species ’.[4]

Such a situation is not unique to taxonomic terms. For example, since some grapefruit have yellow, sour flesh and others have pink, sweet flesh, the fuzzy concept ‘ grapefruit ’ could be replaced by two precise concepts . . . ‘ Florida-grapefruit ’ and ‘ Texas-grapefruit ’. But in most contexts in which the word ‘ grapefruit ’ is used, everything that is said about one kind of grapefruit is true of the other. In other contexts the distinction between the two slightly different fruits might be important and the distinction could be made. However, ‘ grapefruit ’ cannot be defined traditionally so that it is impossible to make a true statement about grapefruit in general which is false with respect to some individual grapefruit.

---

[1] Grant, 1957, pp. 46, 61
[2] Cain, 1954, p. 106
[3] Hull, 1964, pp. 1-11
[4] Simpson, 1961, p. 163

An example which is more like ' species ' is the word ' dishonesty '. There are at least three different criteria for dishonesty—lying, cheating, and stealing. The fulfilment of any one of these criteria is sufficient for proper application of the predicate ' dishonest '. If someone lies, he is being dishonest. If someone cheats, he is being dishonest. If someone steals, he is being dishonest. But it does not follow that if someone performs a dishonest act, he has stolen. He may have lied or cheated or both. Thus, a traditionalist would say that there are three kinds of dishonesty and want to abandon the word ' dishonesty ' for three more precise words. In a sense there are three kinds of dishonesty, but it is not a sense in which the term ' dishonest ' is fuzzy, meaningless, or indefinable. M. J. Scriven says of this example:

> There are cases of special interest here where we are led to apply the same term even when no natural cluster exists; the classic example is that of dishonesty, where we may discover that lying, cheating and stealing, the three main criteria, are not correlated with each other. This does not destroy the social utility of the term, though it may impair its value for personality-theory.[1]

There are several criteria for determining species as evolutionary units. Several criteria are required because evolutionary development is extended in time and space, because numerous factors affect this development, and because there are various ways in which species reproduce themselves. This does not destroy the general utility of the term ' species ' in biology, though it may impair its value in particular areas.

To some the ' solution ' to the species problem presented in this paper may seem no solution at all (which means it's not the kind of solution they had in mind). Instead it might seem merely philosophical sleight of hand. All that is necessary to solve any problem in definition is to trot out a disjunctive definition. Two examples in which this procedure cannot be used may help to dispel this illusion of fakery. In the early days of genetics, genes were held to be the units of heredity. For each gross trait such as eye colour there was supposedly a discrete particle which controlled it. When a radical change occurred in a trait, a gene mutation was postulated as the cause. As genetics developed ' gene ' was also defined operationally as the smallest unit of recombination. It was assumed that the smallest units of heredity, mutation, action, and crossing over were discrete units and one and the same unit. However, as larger and larger numbers were taken into consideration as more conveniently studied organisms were investigated, the limits of what had been envisaged as a gene were passed right through. The progression stopped only at the molecular level with the nucleotide pair (in the case of DNA), and no one wanted to call a pair of nucleotides a ‘gene’ in its original sense.

---

[1] Scriven, 1959, p. 867

In place of the classic gene several new units have been introduced. The smallest unit of mutation is termed the muton and consists of one nucleotide pair. The smallest unit of recombination is termed the recon and is also one nucleotide pair long. The unit of information, the codon, is most probably three nucleotides long. What comes closest to the classic gene is the cistron, the unit of function, which averages about five hundred nucleotide pairs in length. But the product of one cistron is not one trait but one protein. And the whole situation is complicated further by the possible existence of operators which control one or more cistrons, providing a larger natural unit, the operon. In any case nothing approximates the classic gene. This fact presents no serious barrier to the definition of ‘gene’. Scientific words often acquire additional or even radically different meaning as science progresses, but in the case of the word ‘gene’ all but two of the meanings happen to conflict. Almost nothing a geneticist would want to say about a muton or recon would he want to say about a codon, a cistron or an operon, and almost nothing he would want to say about any of the new units would he want to say about the classic gene. The fate of the word ‘gene’ in modern genetics is that except in general discourse or in the first few chapters of an introductory text (and its use here is debatable), its function has entirely evaporated. Using such a truly ambiguous term like ‘gene’ in intermediate or advanced discourse can lead to nothing but confusion.

‘Protoplasm’ is an earlier example. It has already passed out of scientific discourse. Not too many years ago protoplasm was thought to be the stuff of life. If only its complex structure could be analysed, the secret of life would be revealed. As it turned out, it was primarily the nucleic acid constituents (a relatively simple molecule) that were of importance in heredity. Perhaps ‘protoplasm’ could be defined now that there is so much knowledge of the molecular make-up of the cell, but it is doubtful whether any scientist would take the time to try since the concept of protoplasm is no longer of any use in biology. Its original function has been entirely usurped by other concepts.

The fate of ‘gene’ and ‘protoplasm’ has not befallen ‘species’. On appropriate definition it is not ambiguous, and it is still an important concept in biology. Thus, the defining of it disjunctively is both possible and justified. The definition of ' species ' presented in this paper may prove inadequate for the purposes of evolutionary taxonomy, but its formulation will have shown at least in what respects a definition of its kind is superior to the typical Aristotelian definitions of ' species ' thus far offered in the literature.

6 *Conclusion*

It is commonly held that Darwin shattered the essentialist or static concept of species. Darwin shattered something all right, but it was only the first two tenets of essentialism, the ontological and methodological assertions. Darwin did not alter taxonomists' predisposition for Aristotelian definitions, either for species names or for ' species'. Contemporary taxonomists have finally abandoned attempts to define taxa names traditionally, but they still refuse to accept any but an Aristotelian definition of ' species '. To be sure, there are many objections to abandoning the clear-cut analytic-synthetic distinction which the adoption of cluster concept analysis implies, not the least of these being that it has broad philosophical ramifications; but taxonomists have already committed themselves to the necessity of defining taxa names as cluster concepts. No new logical or philosophical obstacles stand in the way of their extending this practice to include ' species '.

Dept. of Philosophy,
University of Wisconsin,
Milwaukee,
Wisconsin

ORIGIN AND CONCEPT OF RELATIVITY (II)*

G. H. KESWANI

5 *What did Poincaré, Lorentz, and Einstein themselves say about the Authorship of the Theory?*

Poincaré lived for about seven years after the publication of Einstein's paper but nowhere did he credit Einstein with the discovery of the theory of relativity. On the contrary, he stated often that Lorentz was the discoverer. Indeed at one place in the year 1912 while mentioning some other contribution of Einstein he spoke at the very same place[1] of ' le principe de relativité de Lorentz ', and he said so repeatedly. And there is a reason for this. Lorentz was the first to use L.T.E. to explain the null results of experiments undertaken to discover the absolute motion of the earth, although he did not explicitly enunciate the principle of relativity, in his memoir of 1904. Indeed, L.T.E. form the vehicle in which the principle of relativity moves and establishes its dominion in the world of physics wherever high velocities are encountered.

Of course, Poincaré was well aware of Einstein's work and in the year 1911 spoke in glowing words about him when Poincaré was asked to give his opinion of Einstein for a certain appointment,[2] but at no place did he attribute the theory of relativity to Einstein. We may mention a corroborative episode narrated by Max Born.[3] In a lecture on relativity given by Poincaré in 1909, at which Born was present, Poincaré did not so much as mention the name of Einstein although he mentioned other names and in fact gave Max Born the impression that he was recording Lorentz's work.

While Poincaré considered Lorentz as the discoverer, Lorentz himself never made any claim to the discovery. However, he sometimes credited Poincaré with the discovery and sometimes Einstein solely. Commenting on the contributions of Poincaré to mathematical physics, he said in the year 1914:

---

* Part I appeared in the previous number.

[1] *Oeuvres de Henri Poincaré*, IX, Gauthier-Villars, Paris, 1954, p. 680, also see pp. 654–655. Poincaré died in 1912. In what follows we refer to this volume as *Oeuvres*.

[2] Carl Seelig, *Albert Einstein*, Staples Press, English trans., 1956, p. 134

[3] Max Born, *Physics in My Generation*, Pergamon Press, 1956, p. 192

> ' Ce furent ces considérations publiées par moi en 1904 qui donnèrent lieu à Poincaré d'écrire son Mémoire sur la dynamique de l'électron, dans lequel il a attaché mon nom à la transformation dont je viens de parler. Je dois remarquer à ce propos que la même transformation se trouve déjà dans un article de M. Voigt publié en 1887 et que je n'ai pas tiré de cet artifice tout le parti possible. En effet, pour certaines des grandeurs physiques qui entrent dans les formules, je n'ai pas indiqué la transformation qui convient le mieux. ' Cela été fait par *Poincaré et ensuite par MM. Einstein et Minkowski*.'[1]

> ' Poincaré, au contraire, a obtenu une invariance parfaite des équations de l'électrodynamique, et il a formulé " le postulat de relativité ", termes qu'il a été le premier à employer.'[2]

We witness a strange spectacle of self-abnegation, one in which each renounces his own claim and credits the other.

But later, on many occasions, Lorentz mentioned Einstein's name only. For example, at a conference on the Michelson-Morley experiment held at Mt Wilson in 1927 where Michelson, Miller and Kennedy were present, Lorentz[3] said:

> ' A transformation of the time was also necessary. So I introduced the conception of local time which is different for different systems of reference which are in motion relative to each other. But I never thought

---

[1] *Oeuvres*, p. 685. Also see p. 689.

[2] *Oeuvres*, p. 687. After quoting this passage approvingly, René Dugas, op. cit., p. 649, remarks, ' Indeed, it must be emphasised that Einstein alone crossed the Rubicon by making a principle of what Poincaré qualified as a postulate '. To this distinction between Tweedledum and Tweedledee we must add the comment that Poincaré had used the word ' principle ' or ' law ' rather more often than the word ' postulate ' when referring to relativity and, more, that in his original paper of 1905, Einstein (p. 38) said, ' We will raise this conjecture (the purport of which will hereafter be called the " Principle of Relativity ") to the status of a *postulate* . . .', obviously making little distinction himself between the words ' principle ' and ' postulate '.

[3] *Astrophys. J.*, 1928, **68**, 385-388. The conference also discussed the theory of the Michelson-Morley experiment and the results to be expected. It is truly surprising that there is no settled opinion yet as to what is to be expected in this experiment; but a null result is, of course, consistent with relativity.

that this had anything to do with real time. This real time for me was still represented by the older classical notion of an absolute time, which is independent of any reference to special frames of coordinates. There existed for me only one true time. I considered my time transformation only as a heuristic working hypothesis. *So, the theory of relativity is really solely Einstein's work.*

'Asked if I consider this contraction as a real one, I should answer "yes". It is as real as any thing we can observe.'

Up to the end (1928) Lorentz was reluctant to give up the conception of absolute space and time. Whittaker[1] recalls this.

We now enter a critical stage in our inquiry. Einstein, in a letter written only two months before his death in July, 1955, has left on record that he knew nothing about Poincaré's work and that he knew of only two papers of Lorentz published much earlier, before he wrote his own paper in 1905. We reproduce this letter[2] below:

> 'There is no doubt, that the special theory of relativity, if we regard its development in retrospect, was ripe for discovery in 1905. Lorentz had already observed that for the analysis of Maxwell's equations the transformations which later were known by his name are essential, and Poincaré had penetrated even deeper into these connections. Concerning myself, I knew only Lorentz's important work of 1895—"La théorie electromagnétique de Maxwell" and "Versuch einer Theorie der elektrischen und optischen Erscheinungen in bewegten Kōrpern"—but not Lorentz's later work, nor the consecutive investigations of Poincaré. In this sense my work of 1905 was independent. The new feature of it was the realization of the fact that the bearing of the Lorentz transformation transcended its connection with Maxwell's equations and was concerned with the nature of space and time in general. A further new result was that the "Lorentz invariance" is a general condition for any physical theory. This was for me of particular importance because I had already previously found that Maxwell's theory did not account for the microstructure of radiation and could, therefore, have no general validity.'

Respectfully, we must submit as follows.

It appears that Einstein was recording from memory. The two papers of Lorentz were not both published in 1895 but in 1892 and

---

[1] E. T. Whittaker, *A History of the Theories of Aether and Electricity* (1900-1926), Thomas Nelson, 1953, p. 36

[2] Max Born, op. cit., p. 194

1895 respectively.[1] None of these papers, of course, contains L.T.E. or relativity. Nor do they contain 'Lorentz's theory of the electrodynamics of moving bodies' in any form that can be regarded as 'in agreement with the principle of relativity' on the basis of Einstein's 'kinematical principles'. The particular reference in Einstein's paper to Lorentz's work quoted earlier cannot be explained without assuming that Einstein was referring to Lorentz's theory propounded in the memoir of 1904. Writing in the year 1905, having published seven papers dealing with problems at the forefront of physics, showing marked awareness of Lorentz's work on the electrodynamics of moving bodies, elaborating what he considered a new and revolutionary theory of the electrodynamics of moving bodies, and considering it necessary to relate his work to that of Lorentz, did Einstein find it enough to have studied only those papers of Lorentz which were published many years ago?

As to Poincaré, the evidence is conclusive that Einstein had studied his *Science and Hypothesis*, in which the principle of relativity was first elaborated.

We must again state that Lorentz did not posit L.T.E. only to establish invariance of Maxwell's equations—indeed he did not establish complete invariance of Maxwell's equations. Lorentz too assumed that in spite of L.T.E. (instead of Galilean equations of transformation), the laws of mechanics were the same in different systems. He said (p. 27), 'It is clear that the state we have supposed to exist in the moving system will really be possible, if in $\sum$ and $\sum'$, the products of the mass m and the acceleration of an electron are to each other in the same relation as the forces. . . .'

We must also submit that there is no reference to 'Lorentz-invariance', i.e. invariance of the laws of physics under the new transformation equations in Einstein's paper of 1905. As we pointed out earlier it was Poincaré who used the words 'invariants of Lorentz group' and related ideas for the first time in a paper published in January 1906. (We are, of course, not insisting on the use of the *word* Lorentz in this context.) However, Einstein (p. 51) did mention that such 'parallel' transformations form a group.

It is further said that since Maxwell's equations cannot account for micro-phenomena, the notion of Lorentz-invariance is more general. This is certainly true but there is no statement to this effect in Einstein's original paper. Einstein himself used invariance of a *macro-phenomenon*, the principle of constancy of velocity of light, to derive L.T.E.[1]

---

[1] H. A. Lorentz, *Collected Papers*, Martinus Nijhoff, The Hague, 1936, 2, 164, and 5, respectively

6 *Relativity according to Poincaré, Lorentz, and Einstein*

Firstly regarding Poincaré's relativity. Poincaré used the phrase 'the law of relativity' initially in the context of relativity of space,[2] i.e. equivalence of all points in space for description of the laws of physics ('local' effects apart). This doctrine, which has received much less attention than the relativity of motion, is perhaps equally important. How does it come about that all points in space are equivalent?[3] No experiment within the solar system, which as a whole is in motion relative to the system of the stars, has indicated that there are differences in the laws of physics at different locations in space or for different directions. Is it that in order 'to apply the law of relativity in all its rigour, it must be applied to the entire universe', as Poincaré[4] conjectured? We shall return to this question in a later paper.

As another step towards generalisation, Poincaré used the term 'the principle of relativity' to include relativity of space as well as of uniform motion. Poincaré, however, saw immediately that the principle

[1] In fact this principle, postulating invariance of the geometry of the wave-surface of electromagnetic propagation (the second postulate), is a particular case of the invariance of the wave-equation which in turn is a particular case of invariance of Maxwell's equations.

It can be shown easily that the assumption of invariance of Maxwell's equations leads to L.T.E. (More accurately, the assumption of invariance of Maxwell's equations leads to a wider group, namely the conformal group. Further assumptions (e.g. linearity) are necessary to extract the Lorentz group. We are indebted to one of the referees for this remark.) In fact, a much weaker condition of invariance of the velocity of an electromagnetic pulse, coupled with plausible assumptions based on the physical situation, is *sufficient* to establish L.T.E., as we shall see in a subsequent paper.

An interesting question is: To what extent do Maxwell's equations already imply their own invariance? Is it, as Max Born conjectures, that Lorentz transformation is an intrinsic property of Maxwell's equations? (*Natural Philosophy of Cause and Chance*, Oxford, 1949, p. 27).

A superficial examination of Maxwell's equations at once shows that charge is already implied to be invariant, i.e. its measure (ρ) is independent of its motion (u), because we write

$$\operatorname{curl} H = \frac{1}{c}\left(\frac{\partial E}{\partial t} + \rho u\right),$$

for all values of u.

[2] H. Poincaré, *Science and Hypothesis*, Dover, p. 77

[3] Einstein (p. 44) uses the words 'homogeneity which we attribute to space and time'.

[4] H. Poincaré, op. cit., p. 77

was not altogether of general validity and, in particular, that it was not true of accelerated motion or even of *uniform* rotation. If the sky were for ever covered with clouds, said Poincaré, and if we had no means of observing the stars, we might, nevertheless, conclude that the earth turns round, because flattening at the poles and Foucault's pendulum would reveal this rotational motion.[1] No observations on the stars need be carried out to determine the motion of the earth with reference to the absolute system of the stars. (We shall see in a later paper how the possibility that it is the system of the stars that rotates and not the earth must be excluded.) The rotational motion is, therefore, not relative to any reference frame, but relative only to one system and one only; it is absolute. The principle of relativity is not, therefore, generally valid. That is why Poincaré thought it was ' battered '. Pondering over the effects of rotational motion, he asked, 'Can a thing turn without turning with respect to something? '[2] And what could be this something, which is, as it were, everywhere?

Rather fortunately, the young Einstein was not beset with these doubts about the principle when applied to non-uniform motion and launched relativity on its triumphant career that we know of so well. Of course, he had to return to Poincaré's misgivings later, but with the courage only a youth could possess, Einstein saw only what need have been seen then (1905), a partial but inspiring vision. Poincaré saw it complete with long shadows of doubt!

Nevertheless, Poincaré continued to develop the principle of relativity of uniform motion of translation and obtained almost all the essential results before Einstein. Occasionally he paid lip-service to the aether [3] which he never defined sufficiently clearly or, at any rate, categorically. There is not the slightest doubt that Einstein (p. 38) was the first to proclaim unequivocally that the aether-concept is superfluous if the principle of relativity is to hold (not to mention important developments such as the formula for compounding velocities and the relativistic theory of the Doppler effect). But he was able to do this because he disregarded (or probably did not appreciate) the difficulties the principle of relativity had then (1905) to face if non-uniform and rotational motions were considered.

[1] H. Poincaré, op. cit. p. 114

[2] Ibid., p. 114

[3] In the discussion on pages 169-172, ibid., for example, Poincaré categorically denies the aether, but the remark at the top of page 244, ibid., makes his position uncertain. This state of belief persisted right up to his death.

Whereas Poincaré and Einstein in their ideas on relativity meant more nearly the same thing, Lorentz's basic ideas were different, as he himself told us again and again. He firmly believed in an aether, the domain of absolute time and absolute space. Motion relative to this aether produced 'real' changes. In 1921 he compared[1] the change in length due to motion, to the change produced by temperature. Ehrenfest,[2] who succeeded Lorentz in his chair at Leiden, tells us clearly about Lorentz's conception of the aether:

> 'Stationary aether meant for Lorentz, just as it already did for Frenesl, stationary with respect to something like the world of the fixed stars.'

What are the physical consequences of postulation of such an aether? Are the laws of physics within various systems moving uniformly with reference to this aether different? If not, what is the use of postulating an additional reference frame of the aether? To these and other questions, we shall turn in subsequent papers. It is to be clearly noted, however, that aether *can* consistently be postulated if it is an inertial system and experimental results consistent with relativity will not thus be violated. The question really is whether it is necessary to do so and if so what are the physical consequences of such an aether.[3]

7 *Minkowski*

Although Einstein rejected the notion of the aether or an absolute frame of reference, derived L.T.E., which had only been posited ad hoc by Lorentz earlier, and established some new results in his paper of 1905, there is evidence that he had not yet fully dominated the theory he propounded. But before we critically comment on Einstein's paper, we must mention a revolution which turned the theory from its somewhat turbid beginnings into clarity in physical interpretation and mathematical representation alike.

---

[1] H. A. Lorentz, *Nature*, 1921, **106**, 794

[2] *Collected Scientific Papers of Paul Ehrenfest*, North-Holland Pub. Co., 1959, p. 474

[3] Years ago (in 1911) von Laue said, 'A really experimental decision between the theory of Lorentz and the theory of relativity is indeed not to be gained, and that the former, in spite of this, has receded into the background, is chiefly due to the fact, that close as it comes to the theory of relativity, it still lacks the great universal principle, the possession of which lends the theory of relativity . ... an imposing appearance.' Quoted in Max Jammer, *Concepts of Space*, Harvard Univ. Press, 1954, p. 142. H. Margenau and R. A. Mould more recently wrote similarly, 'If the laws of nature are of the same form with respect to all frames of reference, then they are of that form in any chosen frame of reference. . . . Clearly it does not make sense to speak of settling the matter between Einstein and Lorentz by experiment. . . . What is significant about Einstein's principle is that it *can* be maintained, not that it must be maintained.' 'Relativity: An Epistemological Appraisal', *Philosophy of Science*, 1957, **24**, 303.

The man who brought this about was Minkowski. He went for his inspiration neither to Lorentz nor to Einstein, though he closely traced and extended Poincaré's line of thought. Minkowski was likely not aware of Poincaré's work. His twin conception of a 4-dimensional space-time continuum and its invariant metric stemmed from purely geometrical roots. Indeed it was he who introduced tensors in the theory of relativity. So revolutionary was Minkowski's way of presentation and analysis that Einstein[1] remarked, 'Since the mathematicians have invaded the theory of relativity, I do not understand it myself any more'. Minkowski was the first true mathematician of relativity.

What is remarkable is that according to Minkowski's way of presentation 'relativity' became less significant. The keyword here was 'invariants', which remained the same no matter what the frame of reference in which they were measured. The simplest of these was the 4-dimensional metric of space-time. A particle in uniform translational motion describes an invariant 4-dimensional 'word-line', an element of which at the 'world-point' (x, y, z, t in *any* inertial frame) is

$$d\tau = \frac{1}{c}\sqrt{(c^2dt^2 - dx^2 - dy^2 - dz^2)},$$

being physically the (proper) time measured by a clock stationary at the particle itself. Whatever be the frame of reference, $d\tau$ remains the same. It is to be noted, however, that $d\tau$ is not a complete differential. Thus, for two world points *connected by two different world lines*, $\int d\tau$, the integral of 'proper' time, does not have the same value. However, the world lines of two particles in *uniform* motion can intersect in no more than one point.

Minkowski drew attention to the *absolute elements*, so to speak, preserved in the flux of space-time (world) which did not change no matter what the frame of reference might be. Indeed, without Minkowski's ideas, the development of the general theory of relativity would not have been possible, as Einstein himself acknowledged later.

---

[1] Quoted by A. Sommerfeld in *Albert Einstein Philosopher-Scientist*, The Library of Living Philosophers, 1949, p. 102

It is interesting to know what Minkowski himself thought of the origin of the ideas associated with relativity-theory. He has left on record [1]:

> 'The validity without exception of the world-postulate, I like to think, is the true nucleus of an electromagnetic image of the world, which *discovered by Lorentz*, and *further revealed by Einstein*, now lies open in the full light of day.'

8 *Criticism of Einstein's Paper*

As we remarked earlier, Einstein was the first to *derive* L.T.E. from the postulates of (i) relativity and (ii) constancy of velocity of light in all frames of reference. Lorentz had only assumed these equations, and though Poincaré clearly saw that these equations were consistent with the principle of relativity, and apparently also saw that they demanded an upper limit of velocities equal to the velocity of light, he made no attempt to derive L.T.E. Did he consider this trivial? We do not know.

It appears to us, however, that Einstein's original derivation is not satisfactory. Einstein (pp. 44-45), uses velocities of light propagation equal to $(c-v)$, $(c+v)$ and $\sqrt{(c^2-v^2)}$ in utter disregard of his own second postulate. We must note that these velocities of light-propagation, in Einstein's derivation, are physically measured (i.e. observable) velocities and not certain quantities in an algebraic calculation. It is interesting to see exactly how Einstein established L.T.E. and we, therefore, give his method along with some comment in Appendix A.[2]

Nor was Einstein yet in possession of the 4-dimensional concept. After developing Lorentz's equations, Einstein (p. 46) uses them to transform the spherical wave-front of light in one frame given by the equation

$$x^2+y^2+z^2-c^2t^2=0$$

into another frame moving relative to the first and gets

$$x'^2+y'^2+z'^2-c^2t'^2=0.$$

He (p. 46) rightly concludes:

> 'The wave under consideration is, therefore, no less a spherical wave with velocity of propagation c when viewed in the moving system.'

But when considering the energy contained within the expanding wave-front of *light* as observed in the two systems, Einstein (p. 57) asserts:

---

[1] H. Minkowski, 'Space and Time', *The Principle of Relativity*, Dover Publications, p. 91

[2] Also see the latter part of footnote 1, p. 294, Part-I.

> 'The spherical surface—viewed in the moving system—is an ellipsoidal surface. . . .'

He then proceeds to work out the equation of the expanding ellipsoid!

9 *Concluding Remarks*

The one conclusion to be drawn is that each one of us stands on the shoulders of others and only thus can even a genius see farther than others. The development of relativity-theory also conforms to this.

It was Poincaré, first and foremost, who talked of relativity of displacements or space and of the principle of relative motion for uniform translation, and of various novel ideas associated with the theory of relativity. However, he noticed that the principle was not true of non-uniform motion. It was possible to find absolute effects due to non-uniform motion. Poincaré did not, therefore, categorically reject the concept of the aether or an absolute frame of reference. Although he was consequently beset with doubts, he continued to develop the principle of uniform relative motion and it was he who recognised the connection between the principle of relativity and Lorentz transformation equations, posited earlier by Lorentz.

Einstein knew of the pioneering ideas of Poincaré and probably Lorentz's transformation equations, but unfortunately he did not say in his original paper what was strictly his own contribution; there is no reference to any previous work at all excepting a general reference to Lorentz's theory of electrodynamics. But Einstein did not only advance the subject analytically and derive L.T.E.—not in a satisfactory way, though—but he was also the first to reject the aether-concept unequivocally, thus becoming the first true believer in the creed of relativity, its prophet. The prophet had, however, apparently paid no heed to the uncertainties expressed by Poincaré the doubter, then. It was a blessing that he did not.

To measure the relative contributions to the principle of relativity, we might quote Max Born[1]:

> 'Einstein's work was the keystone to an arch which Lorentz, Poincaré and others had built and which was to carry the structure erected by Minkowski. I think it is wrong to forget these other men. . . .'

and we might add what every mason knows; the keystone could be placed only when the rest of the arch was there and the falsework of the aether could be removed only after the whole arch was complete.

---

[1] Max Born, *Physics in My Generation*, Pergamon Press, 1956, p. 195

But that is not the only story. The problem of non-uniform motion had yet to be tackled and the very principle of relativity subjected to a more critical examination, as we shall see. All things must change and man must continue

> . . . . to hope till hope creates
> From its own wreck the thing it contemplates.
>
> P. B. Shelley, *Prometheus Unbound*

B18, South Extension (Part II),
New Delhi, 16
India

APPENDIX A

Einstein's Original Derivation of L.T.E.

We first reproduce Einstein's demonstration and then briefly comment on it. (*The Principle of Relativity*, Dover, pp. 43-46).

> To any system of values x, y, z, t, which completely defines the place and time of an event in the stationary system [K], there belongs a system of values ξ, η, ζ, τ, determining that event relatively to the [moving] system k, and our task is now to find the system of equations connecting these quantities.
>
> In the first place it is clear that the equations must be *linear* on account of the properties of homogeneity which we attribute to space and time.
>
> If we place $x' = x - vt$, it is clear that a point at rest in the system k must have a system of values x', y, z, independent of time. We first define τ as a function of x', y, z and t. To do this we have to express in equations that τ is nothing else than the summary of the data of clocks at rest in system k, which have been synchronized according to the rule given in § 1.
>
> From the origin of system k let a ray be emitted at the time τ_0 along the X-axis to x', and at the time τ_1 be reflected thence to the origin of the co-ordinates, arriving there at the time τ_2; we then must have $\frac{1}{2}(\tau_0 + \tau_2) = \tau_1$, or, by inserting the arguments of the function τ and applying the principle of the constancy of the velocity of light in the stationary system:
>
> $$\tfrac{1}{2}\left[\tau(0, 0, 0, t) + \tau\left(0, 0, 0, t + \frac{x'}{c-v} + \frac{x'}{c+v}\right)\right] = \tau\left(x', 0, 0, t + \frac{x'}{c-v}\right).$$
>
> Hence, if x' be chosen infinitesimally small,
>
> $$\tfrac{1}{2}\left(\frac{1}{c-v} + \frac{1}{c+v}\right)\frac{\partial \tau}{\partial t} = \frac{\partial \tau}{\partial x'} + \frac{1}{c-v}\frac{\partial \tau}{\partial t},$$

or
$$\frac{\partial\tau}{\partial x'}+\frac{v}{c^2-v^2}\frac{\partial\tau}{\partial t}=0.$$

It is to be noted that instead of the origin of the co-ordinates we might have chosen any other point for the point of origin of the ray, and the equation just obtained is therefore valid for all values of x', y, z.

An analogous consideration—applied to the axes of Y and Z—it being borne in mind that light is always propagated along these axes, when viewed from the stationary system, with the velocity $\sqrt{(c^2-v^2)}$, gives us

$$\frac{\partial\tau}{\partial y}=0,\ \frac{\partial\tau}{\partial z}=0.$$

Since τ is a *linear* function, it follows from these equations that

$$\tau=a\left(t-\frac{v}{c^2-v^2}x'\right)$$

where a is a function $\phi(v)$ at present unknown, and where for brevity it is assumed that at the origin of k, $\tau=0$, when $t=0$.

With the help of this result we easily determine the quantities ξ, η, ζ by expressing in equations that light (as required by the principle of the constancy of the velocity of light, in combination with the principle of relativity) is also propagated with velocity c when measured in the moving system. For a ray of light emitted at the time $\tau=0$ in the direction of the increasing ξ

$$\xi=c\tau \text{ or } \xi=ac\left(t-\frac{v}{c^2-v^2}x'\right).$$

But the ray moves relatively to the initial point of k, when measured in the stationary system, with the velocity $c-v$, so that

$$\frac{x'}{c-v}=t.$$

If we insert this value of t in the equation for ξ, we obtain

$$\xi=a\frac{c^2}{c^2-v^2}x'.$$

In an analogous manner we find, by considering rays moving along the two other axes, that

$$\eta=c\tau=ac\left(t-\frac{v}{c^2-v^2}x'\right)$$

when

$$\frac{y}{\sqrt{(c^2-v^2)}}=t,\ x'=0.$$

Thus

$$\eta = a\frac{c}{\sqrt{(c^2-v^2)}}y \text{ and } \zeta = a\frac{c}{\sqrt{(c^2-v^2)}}z.$$

Substituting for x' its value, we obtain

$$\begin{aligned}\tau &= \phi(v)\beta(t-vx/c^2),\\ \xi &= \phi(v)\beta(x-vt),\\ \eta &= \phi(v)y,\\ \zeta &= \phi(v)z,\end{aligned}$$

where

$$\beta = \frac{1}{\sqrt{(1-v^2/c^2)}},$$

and ϕ is an as yet unknown function of v.'

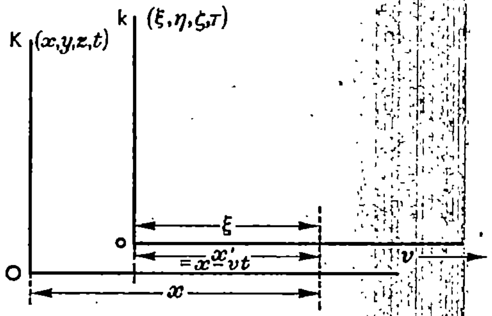
$\phi(v)$ is later shown to be equal to unity. Incidentally, there is an inconsequential notational slip. The function a above is also called $\phi(v)$ but in fact

$$\phi(v) = \frac{ac}{\sqrt{(c^2-v^2)}} = 1$$

$$\text{i.e. } a = \sqrt{(1-v^2/c^2)}.$$

As we mentioned earlier, Einstein used velocities of light-propagation equal to $(c-v)$, $(c+v)$ and $\sqrt{(c^2-v^2)}$ overtly, and yet got the correct end result, the usual expressions for L.T.E. The steps taken have a curiously compensating effect and apparently the demonstration was driven towards the result.

The annexed figure, we hope, corresponds to the above demonstration. In this comment we consider happenings along the x and ξ axes only.



Einstein introduces the moving co-ordinate $x' = x - vt$ so that: 'a point at rest in the system K must have a system of values x', y, z independent of time'.

A ray leaves the origin o of K at time τ_0, travels a distance ξ, as measured in K, in time $(\tau_1-\tau_2)$ and returns to the origin o of K at time τ_2. Applying the principle of constancy of the velocity of light in the moving system K, Einstein writes the correct relationship:

$$\tau_1 = \tfrac{1}{2}(\tau_0+\tau_2).$$

However, the times of departure or arrival corresponding to the three events as measured in the stationary system K are given as follows, again, it is asserted, 'applying the principle of constancy of the velocity of light in the stationary system':

$$t,\ t+\frac{x'}{c-v} \text{ and } t+\frac{x'}{c-v}+\frac{x'}{c+v}.$$

The statement, 'the ray moves relatively to the initial point (origin) of K when measured in the stationary system (K), with the velocity $(c-v)$, etc.' is physically difficult to follow. We cannot meaningfully speak of velocity of a light-ray relative to a point of *one* inertial system (the origin of K) but as it (velocity) is measured in *another* system (K). The argument seems to proceed as follows: x' is *some* length (measured in the stationary system K) which the outgoing and returning rays traverse with velocity $(c-v)$ and $(c+v)$ relative to the origin of the moving system K, the quotients $\frac{x'}{c-v}$ and $\frac{x'}{c+v}$ being the corresponding time intervals, as measured in K. The procedure is operationally vague and certainly not in conformity with the second postulate.

It is to be noted that $(c-v)$, $(c+v)$ and $\sqrt{(c^2-v^2)}$ are, in Einstein's exposition, considered as velocities of light-rays in K and not as certain quantities in an algebraic process.

Incidentally, although Einstein used the word 'event', he did not use the 4-dimensional space-time concept. For a similar remark see: N. H. de V. Heathcote, *Nobel Prize Winners in Physics* (Schuman, New York, 1953, p. 195). The latter concept was due to Poincaré (1906), as we remarked earlier in this paper, and was further developed, apparently independently, by Minkowski. In 4-dimensional space-time, the fourth co-ordinate is *not* time but $\sqrt{-1}\times$time.

CAN A MACHINE BE CONSCIOUS?[1] *

DENNIS THOMPSON

'CAN machines think?' is a question which in various forms has been the subject of a great deal of speculation in recent years. Yet the question itself has not been precisely stated. Nor have the issues relevant to the problem it raises been carefully sorted out. As a result, many of the philosophical implications that are supposed to follow from an answer to the question do not follow at all. It is my aim here to indicate just what issues are, and are not, involved in the machine problem, and what consequences for the philosophy of mind follow if the problem is solved. I shall call the view that machines *can* be conscious the 'Machine Theory', and refer to the claim that machines *cannot* be conscious as the 'Anti-Machine Theory'.[2] I shall indicate why, contrary to assumptions implicit especially in the writings of the Anti-Machine Theorists, the claim that machines *can* be conscious implies comparatively little for mind-body theories, while the claim that machines *cannot* be conscious does have consequences which most Anti-Machine Theorists would find unpalatable.

First, it is necessary to reformulate the question 'Can machines think?' The fundamental issues at stake in the problem can be brought out more clearly by replacing the term 'think' with the term 'be conscious'. For if we were to demonstrate that a machine could solve a logical problem, recognise dark clouds as a sign of rain, interpret a piece of music or do any other type of thinking, the Anti-Machine

---

* Received 29.v.64

[1] I am grateful to Dr John Lachs and Professor A. J. Ayer for their helpful comments and criticisms.

[2] Some Anti-Machine Theorists are: J. Cohen, 'Can There Be Artificial Minds?', *Analysis*, 1955, **16**, 36–41; G. Jefferson, 'The Mind of Mechanical Man', *British Medical Journal*, 1949; R. Lachman, 'Machines, Brains, and Models', in *Dimensions of Mind*, ed. by S. Hook, New York, 1960; W. Mays, 'Can Machines Think?', *Philosophy*, 1952, **27**, 148–162; G. Negley, 'Cybernetics and Theories of Mind', *Journal of Philosophy*, 1951, **48**, 574–582; K. Popper, *Conjectures and Refutations*, London, 1963, pp. 296–297; A. L. Samuel, 'Some Moral and Technical Consequences of Automation—A Refutation', *Science*, 1960, **131**, 741–742; M. Taube, *Computers and Common Sense*, New York, 1961; P. Weiss, 'Love in a Machine Age', in *Dimensions of Mind*; R. J. Spilsbury, 'Symposium: Mentality in Machines', *Aristotelian Society Proceedings*, 1952 (Supp.) **26**, 1–86.

Theorist would still ask: ' But can machines *really* think—I mean the way *I* do? ' The point of the objection would be that 'thinking' predicates when applied to machines must be used only analogically; human thinking, it would be argued, has a special property which machine ' thinking ' does not have, viz. consciousness. Thus the question ' Can machines think? ' is appropriately reduced to ' Can a machine be conscious? ' In this reformulation there are several possible sources of confusion which must be cleared up.

One possible snare in the question is the term 'conscious '. Since I shall argue that the machine problem becomes clear only if we focus on hypothetical machines which can duplicate all human behaviour, a behaviouristic description of consciousness would allow the question to be decided *ex hypothesi*. A more fruitful approach is to describe consciousness in the way that would be most favourable to the Anti-Machine Theorists. Most of these theorists apparently regard consciousness as somehow a defining attribute of all ' inner ' experience. One way of describing what is meant by consciousness in this sense might be to say that it is the total ' inner ' difference between being awake and being in a dreamless sleep. Ultimately, however, all that the Anti-Machine Theorist is able to do is to appeal to the reader to examine introspectively his conscious experience. But even on a non-behaviouristic analysis of consciousness, the Machine Theory does not have the consequences that the Anti-Machine Theorists fear. Acceptance or rejection of the Machine Theory does not entail any particular view of consciousness.

The other principal term in the question ' Can a machine be conscious? ' is of course ' machine ', which like many terms has no precise general definition since there is no one quality common to all machines. The machine which is most relevant to the question is the *robot*, which may be described as a mobile machine, outwardly resembling a human, controlled by a computer unit (the programme of which is independent of subsequent human interference) and capable of performing anything a human can perform.[1] The point of the terms ' can ' and ' a ' in the question is to focus attention on such machines not yet built.

Here one of the most common and serious mistakes which arise in discussions of the machine problem should be mentioned. Too often the argument about whether machines are conscious degenerates into a technological treasure hunt, which the philosopher would be wise to avoid.[1] If he is interested in the philosophical implications of the Machine Theory, he should assume that it will be possible to construct the sort of machine which some cyberneticians claim to be possible (e.g. a robot), and study the problem from there. (This is not to deny that there may be philosophical problems at various stages in the technological evolution of robots.)

---

[1] A fictional portrayal of the sort of machine I have in mind is to be found in Karel Čapek's play, *R.U.R.* However, it is irrelevant to my argument whether a 'robot' is an artefact (i.e. made by man for a purpose) or a natural organism (e.g. a product of evolution).

The terms in 'Can a machine be conscious?' should be sufficiently clear now. Anti-Machine Theorists find three barriers to giving an affirmative answer to this question: (1) uncertainty as to what characteristics machines would need to possess in common with humans to be justifiably called conscious; (2) puzzlement about the meaning of 'conscious'; and (3) fear of the philosophical consequences that are thought to follow from the Machine Theory.

The first barrier to acceptance of the Machine Theory should not be too difficult to overcome. It should be clear, irrespective of one's metaphysics, that the only way of deciding whether another person is conscious is by observing his behaviour and his bodily state. If performance is sufficient justification for calling humans conscious, then it is surely sufficient justification for calling machines conscious—even if consciousness is regarded as a property of some sort of 'inner' life. Indeed, this step of the argument is implicitly accepted by Anti-Machine Theorists. This is why the machine debate so often centres around a comparison of what machines and humans can do—the *technological* question.[2]

---

[1] The danger for the philosopher in this sort of discussion is illustrated by a book by Mortimer Taube, *Computers and Common Sense*. Taube charges that computer enthusiasts continually mention 'a great many names of non-existent machines whose operations are described and debated as though they were real'. In a review of Taube's book in *Science* (2nd March, 1962), Walter Reitman mentions some *actual* computers whose programmes do what Taube asserts no computers can do. Taube asserts that Simon, Shaw, and Newell's 'General Problem Solver' cannot be built. Reitman replies that he himself has worked with the 'General Problem Solver' and 'can testify that it does in fact exist, even as you and I'.

[2] Machine Theorists have shown that machines can behave adaptively, learn by trial-and-error, show evidence of self-recognition, make mistakes as humans do, exhibit primitive forms of social behaviour, become neurotic, beat their designers in chess, discriminate universal forms and relations, and compose music that is judged by critics to be in some respects better than that produced by human composers. See W. R. Ashby, *Design for a Brain*, New York, 1960; F. H. George, *The Brain as a Computer*, New York, 1961; 'The Machine Closes In', *Time* (February, 1962), 65;

The technological question usually leads to another objection which Anti-Machine Theorists raise—probably the most common argument in the machine debate. Machines cannot be conscious because no machine can truly create anything or do anything *on its own*, but only what is programmed into it. If all that is meant by this objection is that everything the machine turns out is a result of its structure, its programme and stimulation, and the changes it makes in its programme, then the statement is true enough, but it does not strengthen the Anti-Machine Theory. For everything a human produces is a result of his structure, the stimuli he receives, and what he does with that he has learned. If the objection is supposed to mean that the machine cannot do anything original because we as humans know everything we programme into it, it is false. It has been often pointed out that computers, because of their high speed operations, are able to give results which are in practice unpredictable.[1] Even W. Grey Walter's simple *Machina speculatrix* has produced patterns of behaviour that were not predictable even though the designer had complete knowledge of the parts of the machine and how they were put together.[2]

The second barrier to accepting the Machine Theory consists of linguistic confusions. The Anti-Machine Theorist admits that behaviour and performance generally are necessary conditions for our calling anything conscious, but maintains that they are not sufficient conditions. It still makes sense, these theorists argue, to ask if a robot, which displays all the behavioural characteristics of consciousness, is really conscious. The reason it makes sense is that, in this case, the question expresses puzzlement about the meaning of ' conscious ', rather than a doubt which further empirical evidence of machines' abilities could settle. In this sense, ' Can a machine be conscious? ' comes to mean ' Can the meanings of the terms " machine " and " conscious " be compatible? '

Incompatibilities may break out anywhere among the meanings of the terms that make up the network of concepts related to ' machine '

D. M. MacKay, ' In Search of Basic Symbols', *Cybernetics—Trans. Eighth Conference*, ed. H. von Foerster, New York, 1952; G. Miller *et al.*, *Plans and Structure of Behaviour*, New York, 1960; C. Shannon, ' A Chess-Playing Machine ', *Scientific American* 1950, 48–51; W. Sluckin, *Minds and Machines*, London, 1960; W. G. Walter, *The Living Brain*, New York, 1953.

[1] E.g. N. Wiener, ' Some Moral and Technical Consequences of Automation ', *Science* 1960, **131**, 1355–1358

[2] Walter, *The Living Brain*, p. 130

and 'conscious'—in particular, 'living', 'constructed by man', 'made out of non-living parts', and their contradictories. There is no need here to examine these incompatibilities in detail. For even if the claim that it is logically impossible for machines to be conscious is accepted, it is possible to show that it does not support the Anti-Machine Theory in the way that it is supposed to.

If 'machine' is now in most contexts incompatible with 'conscious', it is because machines until recently have failed to display enough of the characteristics of human behaviour to justify our regarding them as conscious. 'Machine' once was used to refer only to relatively primitive devices, such as pulleys and windmills, which no one considered calling conscious. As technology advances, 'machine' is used to refer also to more and more complex devices, and genuine disputes about the consciousness of machines arise. It is a mistake then to suppose that we can determine whether a machine *can* be conscious by reference only to linguistic usage. It is true that at any given point in technological development, reference to linguistic usage is decisive in determining whether the Machine Theory is true. It is of course also true that if 'machine' is incompatible with 'conscious' (or if any of the other incompatibilities alluded to above arise), then it is logically impossible for a machine to be conscious. But this impossibility is based on the present state of technology which has influenced the current usages of 'machine' and 'conscious'. Such a prohibition is a shaky one on which to base a philosophical theory.

Therefore to argue that the ascription of consciousness to a machine results in a logical absurdity, as the Anti-Machine Theorists do, is to beg the question. Such an argument adjudicates by reference to meanings what ultimately can be decided only by reference to facts. Hence, there is no *inherent* logical contradiction involved in calling a machine conscious. Since observation of behaviour is the only way of deciding whether other people are conscious, there is no good reason to suppose the same criterion should not apply to machines.

If we were to convince the Anti-Machine Theorists of all that has been said thus far, they would probably still hesitate to call a robot conscious. The Theorists' reluctance raises the third barrier and reveals what may be the strongest motive for their denunciation of the Machine Theory. If the Anti-Machine Theorists ascribe consciousness to machines, they believe they have committed themselves to a full-blown behaviourism or materialism. (Indeed, Machine Theorists, too, sometimes seem to be motivated to argue for consciousness in machines because they believe their theory *does* entail some sort of behaviourism or materialism.)

Acceptance of the Machine Theory does not commit us to any particular theory of the relation of mind and body. In fact, acceptance of the Theory has fewer implications for mind-body theories than does rejection of it. I shall first examine the consequences on the supposition that the view that machines can be conscious is true. Next, I shall examine the consequences on the supposition that the view is false. For the purpose of this article, it is possible to classify in five categories all theories of the relation of mental events to physical events.[1] Two of the theories, Idealism and Parallelism, need not detain us. Idealism, which holds that all things are really mental, seems to me to be plainly false. But whether true or false, no decision for or against the Machine Theory will have any force for or against Idealism since the Idealist would classify Machines themselves as mental in some sense. The second theory, Parallelism, asserts that there is a correlation between mental and physical events, and further that there is no causal connection between the two sorts of events. But this theory is no theory at all, even to the extent that the other mind-body theories are. To explain the correlation of the mental and physical (a fact which no mind-body theory denies), Parallelism says merely that the two kinds of events are correlated. Acceptance or rejection of the Machine Theory could not possibly affect a theory which, in effect, explains or clarifies nothing.

We are left with three categories of mind-body theories which are relevant to the machine problem—Interactionism, Epiphenomenalism, and Materialism. Essentially, Interactionism holds that some mental events cause physical events, and some physical events cause mental events. Epiphenomenalism also admits the existence of mental events. It claims that no mental events cause physical events, no mental events cause mental events, and all mental events are caused by physical events. In this category, I include the very similar Identity Theory, such as that propounded by Herbert Feigl.[2] Materialism asserts that there are no mental events. In this category, I wish to include not only the old-fashioned forms of materialism usually referred to by that term, but also philosophical behaviourism and the various related contemporary linguistic theories that seek to clarify the relation of mental and physical words. Admittedly, these classifications are very rough, and some of the proponents of these theories would object to their classmates. However, for the purpose of showing to what degree mind-body theories are affected by acceptance or rejection of the Machine Theory, I believe this rough classification and the implied distinction between mental and physical events will suffice. Now I shall examine the consequences for mind-body theories, first on the supposition that the Machine Theory is true.

---

[1] By 'mental event', I mean to refer to what has been called at various times in the literature of the mind-body problem 'raw feel qualities' or 'direct experience' or 'the given'. This concept remains vague because the nature of privacy, I believe, has yet to be adequately analysed. By using 'event', I do not intend to decide the question whether the 'mental' is best described as a series of events or as states or as processes.

[2] H. Feigl, 'The Mental and the Physical', *Minnesota Studies in the Philosophy of Science*, Minneapolis, 1956, 2, 370–497

First, we can say, contrary to what is sometimes supposed, that the existence of a conscious machine would not be evidence either for or against Materialism. It is true, as far as we can tell by looking at the parts themselves, there is nothing mental about machines. But neither is there anything mental about other people—if we examine only their constitution. There is nothing special about living matter *as such* that justifies our regarding as conscious whatever is composed of it. The relation of the mental to the physical (or of mental concepts to physical concepts), then, is still a problem even if we accept the Machine Theory. Thus, acceptance of the Machine Theory does not require acceptance of Materialism, unless one has previously accepted it as true on other grounds. Conversely, rejection of the Machine Theory does not entail rejection of Materialism unless one has rejected it on other grounds.

The consequences of the Machine Theory for the two remaining mind-body theories, Interactionism and Epiphenomenalism, are not so easily disposed of. Most Anti-Machine Theorists reject all forms of Materialism and tend to favour Interactionism. Accordingly, it is necessary to see what the consequences for these two theories are.

Both theories agree that physical events cause at least some mental events. The theories disagree about the causal efficacy of mental events: (1) Interactionism holds that some mental events cause physical events, while Epiphenomenalism maintains that no mental events cause physical events; (2) Epiphenomenalism further holds that all mental events are caused by physical events.

Prima facie, the existence of a conscious machine would not favour one or the other of the two theories, any more than does the existence of a conscious human. But perhaps it would be possible to decide these points of disagreement between the two theories by performing some experiments on a robot which we would not be able to perform, for one reason or another, on a human being.

A difference in the structures of machines and humans makes it possible, at least in theory, to obtain some experimental evidence which would bear on the decision between Epiphenomenalism and Interactionism. Although these mind-body theories are not scientific theories, scientific experiments which could be performed on machines but not on humans might give factual evidence which one of the mind-body theories could account for better than the other. Because the human brain is a living organism, each part depends for its *existence as a functioning unit* on many other parts. If we cut or remove one part of the brain, the blood flow to another part may be stopped, and the whole organ might *permanently* cease to function. This complex interdependence of the parts of the brain makes it difficult to carry out experiments of the sort necessary to provide evidence for or against mind-body theories. A second difficulty is that moral attitudes generally prevent tampering with the human brain to the extent necessary in such experiments.

Both of these difficulties are mitigated somewhat in the case of a conscious machine. I am not sure just what our moral attitudes would be toward conscious machines, but I should speculate that we would tend to permit somewhat more tampering with a robot's brain than with a human's brain. One reason might be that we would not consider the robot ' one of us '. If our theory of obligation distinguishes human pain from the pain of other conscious beings, then the ' felt ' obligation toward robots would not be so strong as toward our fellow humans. Accordingly, perhaps we would be as likely to use robots for experiments as we now use dogs. But it may be that since robots exhibit as high a degree of consciousness as humans, our obligation will be equally strong, in which case the experimental difficulties will remain much the same. On the other hand, it may be that robots will not feel pain from the same sort of experimenting that humans do. But most of these speculations can be substantiated only when we are actually confronted with robots.

The first difficulty would not be totally eliminated either. It would be possible, if we constructed the robot with the experiment in mind, to make the various tubes and wires accessible to disconnection. And in theory we could disconnect one or more at a time in various parts of the computer unit without rendering permanently inoperable the machine's brain. If we disconnected a part from the computer that caused it to cease functioning completely, we could re-connect it and try removing another part. It is problematical whether we could disconnect enough of the right parts in a computer brain to substantiate a particular mind-body theory before the machine ceased to be able to report its mental events or ceased functioning entirely. Again only actual experiments could determine this.

If a conscious machine, because of its structure and because of our moral attitudes toward it, does provide a slight advantage over humans for performing experiments, what experiments might be performed which would help to decide between Interactionism and Epiphenomenalism? There does not seem to be any experiment that would decide the first point of disagreement between the two theories, viz. whether some mental events cause physical events or whether no mental events cause physical events. For in a functioning machine, as in a living human, every physical event is preceded by at least one other physical event; therefore, it would never be correct to infer that a given physical event could not have been caused by a prior physical event.

It seems more likely that an experiment could be designed to help decide the second point of disagreement between Epiphenomenalism and Interactionism. Suppose we systematically disconnect various sets of tubes in the robot. If certain mental events occur, or cease to occur, each time we perform the same physical operation, we can correctly conclude that these mental events are probably caused by the physical events. The more physical causes we discover, the greater the probability of the empirical generalisation that all mental events have physical causes. The probability that Epiphenomenalism is true, in other words, would be increased. It is true that for every step in the disconnecting operation that failed to turn up a physical cause for a mental event, the probability that all mental events are caused by physical events would decrease, thereby weakening Epiphenomenalism and strengthening Interactionism. But since very few of the disconnecting operations could be performed while a given mental event is occurring, the decrease in probability would no doubt be so slight as to be insignificant. Thus, if the results of the experiment tend to favour either theory, they would favour Epiphenomenalism. Epiphenomenalism, in other words, is not likely to lose, though it may never win.

It should be clear now that the consequences of accepting the Machine Theory as true are not as great as the Anti-Machine Theorists suppose. If one has rejected Materialism in the case of human beings, one has just as good (or just as bad) reason to reject it in the case of conscious machines. Moreover, acceptance of the Machine Theory does not require rejection of Interactionism; nor does it require acceptance of Epiphenomenalism. It does give an edge to Epiphenomenalism to the extent that it allows that theory the possibility of more verification than it has if machines are not counted as conscious. But even this consequence may be repugnant to some Anti-Machine Theorists, who generally wish to regard the mind as more than a ' mere epiphenomenon ' of the brain and, if they can avoid it, will give no advantage, however slight, to Epiphenomenalism. The price of avoiding it is high, however. For this consequence of accepting the Machine Theory is innocuous compared to the consequences of rejecting it.

If we refuse to regard as conscious a robot which can do everything a human can, we (1) strengthen a certain kind of solipsism, and (2) commit ourselves more fully to Epiphenomenalism than we would if we rejected the Machine Theory. By saying that a robot can do everything a human can do, we imply (as I have shown) that the robot has all the characteristics which justify our calling humans conscious. Thus, if we deny that robots are conscious, as the Anti-Machine Theorists would have us do, we have less reason to say that other humans are conscious. The denial of the Machine Theory thus undermines our justification for the belief in the existence of other minds. What is unfortunate about this consequence is that it amounts to a support of a philosophy which no one holds seriously—least of all Anti-Machine Theorists.

But suppose that Anti-Machine Theorists still believe that other people have minds, despite the fact that their justifications for the belief have been weakened. Still to be faced is the second consequence of rejecting the Machine Theory. By rejecting the Machine Theory we are denying that robots have minds, even though they can do everything a human can do. But if a machine which can do everything a human can do does not have a mind, then there is no reason to suppose that the minds of other people have causal efficacy in the physical world, for it follows on this view that we can account fully for other people's behaviour without reference to a mind. Moreover, if the minds of other people have no causal efficacy, then I have little reason to believe my own does. (If my ' willing my arm to move ' causes anything—which is doubtful—it causes a brain event.) This consequence of rejecting the Machine Theory—that minds have no causal efficacy—is what Epiphenomenalism asserts and Interactionism denies. It decides the first point of disagreement between the theories. While we could not decide this disagreement by accepting the Machine Theory, we *do* decide it if we reject the Theory. Of course, we have not thereby shown that all mental events are caused by physical events (the second point of disagreement), but Epiphenomenalism becomes, if the Machine Theory is rejected, more probable by far than Interactionism.

Rejection of the Machine Theory, therefore, turns out to have more drastic consequences for the Anti-Machine Theorists than acceptance of the Theory. Not only does the rejection support a certain kind of solipsism, but it also virtually commits them to Epiphenomenalism. Acceptance of the Machine Theory, it will be remembered, gives not nearly so much support to Epiphenomenalism, nor does it strengthen solipsism. In neither case is Materialism affected.

In this article I have tried to show that refusal to accept the view that a machine *can* be conscious is a result of a number of muddles, the most important of which is a mistake about what philosophical consequences follow from the Machine Theory. Passing in and out of the machine problem are a great number of philosophical issues, but I have had to leave most of these only briefly stated, and certainly not resolved. My chief concerns here have been to eliminate the more troublesome misconceptions concerning the machine problem and to show how and what philosophical problems are related to it. At the same time, I have claimed that machines *can* be conscious, but if I have shown this, I have also shown that it is not such a radical claim after all.

Balliol College,
Oxford

ON THE DUAL NATURE OF LIGHT*

GEORGE MAGYAR

SINCE the particle-wave duality of light was dealt with by quantum mechanics more than thirty years ago[1], it has never ceased to be a fascinating topic for physicists and philosophers of science.

The fact that there appeared to be a paradox in this dual behaviour, causing difficulties of interpretation, of interference with weak light beams, for example, was recognised from the beginning.

Recently the whole problem was re-examined in a sharpened form by Janossy and co-workers[2] who used a Michelson-type interferometer with an arm-length exceeding the length of the wave-train and correlation techniques with a time-resolution of the order of the coherence time (duration of the wave-train).

The paradox as it occurred in their experiment can be outlined briefly as follows. When in a two-beam interferometer we split the wavefront or the amplitude of a quasi-monochromatic light-beam and reunite the two coherent secondary beams, we observe the characteristic bright and dark interference fringes. As the above-mentioned experiments proved, this is valid even for a very weak beam, where on the average less than one photon travels through the interferometer at a time.

On the other hand, when they put two photo-multipliers in the way of the two beams, to count the incidence of individual photons, they did not get any correlation between the two counters above the normal rate of chance coincidences apart from a very small extra coincidence due to the Hanbury Brown and Twiss effect.[3] Consequently, no splitting of the photons occurs: they choose one path or the other. Nevertheless, if in the first experiment we change the position of the mirror in one arm of the interferometer by a quarter of a wavelength, the dark fringes become bright and vice versa: the photons are redistributed. But this means that the photon was influenced in its behaviour by a mirror (or slit), through which it never passed!

* Received 23.i.64

[1] E.g. M. Born, *Optik*, 1933, Berlin, p. 456

[2] L. Janossy and Zs. Naray, *Acta Phys. Hung.*, 1957, 7, 403

[3] R. Hanbury Brown and R. Q. Twiss, *Nature*, 1956, 177, 27

Quantum-mechanically, of course, nothing extraordinary happened. The wave-function characterises the expected behaviour of *individual* photons, so the fringe-like distribution is the same for a single or a million photons. This is very clearly expressed by Dirac: 'each photon then interferes only with itself. Interference between different photons never occurs'.[1]

Although quantum mechanics gives a satisfactory mathematical description and for many this is the ultimate we can ask for, several physicists (e.g. E. Schrödinger, L. de Broglie, D. Bohm, L. Janossy and others) take this as a sign of weakness or incompleteness.

Recently we carried out an experiment[2] which might help to clarify the interpretation. In this experiment we used two real, independent light-sources, instead of the usual single source. This was made possible by the advent of a new powerful coherent light-source: the optical maser. Here we had two light-waves, sinusoidal for sufficiently long time (approximately $\frac{1}{2}$ microsec.) and containing enough photons to produce interference fringes during that time.

Now, if pressed, this experiment can be described, admittedly more cumbersomely, in quantum-mechanical terms.[3] Thus Dirac's statement still holds formally, because, crudely speaking, in the cell of phase-space, where the interference occurs, the photons are indistinguishable.

However, the literal meaning of his statement can be experimentally disproved very simply: just by blocking one beam (in which we *know* that the photons are completely independent from those of the other beam because the coupling of the two sources is excluded), the interference can be stopped. It occurs only when the two independent beams mix freely. In earlier interferometers this step was ambiguous, because the two beams originated in the same source.

So we offer a simpler interpretation to all these experiments.

First of all the duality problem for light should be treated separately from that of matter. Before the advent of quantum theory the two were thought of as completely different entities. Afterwards, when it was demonstrated that both had wave and particle aspects, people had fallen into the opposite extreme of complete identification of these two forms of physical substance. Close as the analogue between photons and particles is, there are, nevertheless, very fundamental differences. The photon is the only particle which is its own antiparticle. It is also the only particle which does not take part in strong or weak interactions but only in electromagnetic ones. Then, for example, consider a dense beam of photons and of electrons: the electrons will never lose their particle character. Also, the speed of light is not an ordinary one: being independent of the speed of the source or of the detector, it is rather a special ' state ' of something; particles which have a rest mass can never reach it. Thus, no wonder when light travelling with this speed is stopped, a new aspect—the photon aspect of light—is created. With increasing wavelength, even the quantum character of absorption and emission of electromagnetic waves becomes unimportant. This happens just when the interaction is no longer with individual electrons or atoms; strong indication that quantum behaviour is created in the interaction.

---

[1] P. A. M. Dirac, *Quantum Mechanics*, 1958, Oxford, p. 9

[2] G. Magyar and L. Mandel, *Nature*, 1963, **198**, 255

[3] L. Mandel, *Phys. Rev.*, 1964, **134**, A10

Finally, carrying the quantum nature of the electromagnetic field (or field generally) to the extreme would mean a return to the principle of ' action at a distance ' which is contradicted by the whole development of physics. The particular nature of the duality problem, of course, could apply to other types of field on the basis of this argument, but it is the electromagnetic field alone, where we can have a ' source-free ' field in sufficient quantity for experimentation. As far as the other fields are concerned, the question is left open.

This uniqueness must be understood as relative only. We know that light can be transformed into other particles and vice versa, but we also know that the special properties noted above then disappear. Or perhaps become latent, which would point towards a deeper unity, on a different basis.

The arguments for separateness already point towards the essence of our interpretation, which is this: light between emission and absorption should be treated as wave phenomenon. We would only speak of photons when light is interacting with matter. That the particle aspect of light is created in this interaction is indicated by the fact that in the modern quantum theory of light[1] the annihilation operator plays a central role. Quantum mechanics describes the absorption of light by matter, where the concept of photon is important, but does not want to know what happens before (' unobservable '). However, in the experiments described earlier, it is easy to work back from the observed fringe separation, on the basis of a simple wave theory, to calculate the widths of the slits, which the waves passed on their way

[1] R. Glauber, *Phys. Rev.*, 1963, 130, 2529

to the screen. Thus the waves carried information. A less trivial application of this property of waves is shown in an ingenious device, originally proposed by Gabor[1] where the hologram of an object is constituted by 'Fourier analysis' of its diffraction and interference pattern and then these waves are synthesised into the appropriate picture at the receiver.

There is an obvious objection to our interpretation: there is no other way of detecting light waves but to stop them in matter and so unavoidably to 'create' photons. What we observe are photons all the time.

This is perfectly true, but we think one should distinguish two cases here. In one case the photons are just intermediaries in the chain of observation: from their information content we make inferences to what happened before. This is not very uncommon in physics. Although the dark and bright patches of a photographic emulsion, for example, are the results of photon-electron interaction, the information we seek on them might have bearing on numerous other macroscopic or microscopic phenomena. We are interested in the photon aspect in such measurements only when looking for the limiting factors of accuracy (noise, etc.).

On the other hand, photon-particle interaction might be of interest in itself, as in elementary particle scattering experiments. One could say that previously the 'collective' aspect, here the 'individual' aspect, comes to the fore.

In the interference experiment with very weak light we are interested in the integrated effect of a large number of photons interacting with our detector: the waves produce a spatial distribution (fringes) and this is recorded via photons. Similar interpretation could be made of the Hanbury Brown-Twiss extra correlation, where certain instantaneous interference effects of waves lead to a special photon time-distribution in the detectors;[2] we cannot go into this in more detail.

In the counting experiment we exclude the possibility of interference between the two waves, thus what we are recording is the individual interaction of light with two independent counters.

Fairly recent evidence for our case is the interaction of light with light, which was first made observable by the application of laser techniques. It is describable in completely classical terms[1] unless the specific nature of the interaction with the mediator electrons becomes important.

---

[1] D. Gabor, *Journ. I.E.E.*, 1946, **93**, III. 429

[2] L. Janossy, *Nuovo Cim.*, 1957, **6**, 111

The important point is that photons travel only as a potentiality in the light waves. Quantum mechanics describes the (observable) interaction where this potentiality is realised. To ascribe the interference fringe-system to a wave-function, representing a single photon, might be plausible mathematically, but its observation experimentally is impossible for obvious technical reasons.

In a recent paper in this Journal,[2] Agassi showed that the assignment of the micro-world to quantum theory and the macro-world to classical theory is problematical. Here we see the complex nature of this problem for light. Most of the recently discovered effects in optics (intensity correlation, beats between independent beams, etc.) are connected with fluctuations of light beams, where Planck's constant cannot be considered negligible. Hence, belonging to the micro-world, they would call for the application of quantum theory. Nevertheless they are equally well described by the classical or semi-classical[3] theory, sometimes in a much simpler way. These aspects of micro-phenomena we called ' collective ' in contrast to the problem of the mechanism of individual electron-photon interaction, which only quantum mechanics can deal with adequately. The complete equivalence of the semi-classical and quantum theories in this field has been proved recently.[4]

This state of affairs could be regarded as a certain modification of the correspondence principle. Similarly, for light the complementarity principle might be interpreted as the relationship between these two aspects. Here we mention in passing that the generally held view of complementarity as a principle forbidding the simultaneous observation of wave and particle aspects of light is an oversimplification. Vavilov reported[5] the observation of random fluctuations due to individual photons in the bright fringes of an interference experiment, carried out with weak light. However, in such cases neither of the two aspects show up sharply.

We do not exclude, of course, the possibility that the ' wave behaviour ' of photons between emission and adsorption is the result of a complicated interaction with an underlying sub-quantum level. This, if found, would not invalidate our interpretation, just as we do not speak about movements of air molecules when discussing the unusually low temperature of last winter, although we are familiar with Maxwell's Kinetic Theory of Gases.

---

[1] Cf. D. A. Kleinman, *Phys. Rev.*, 1962, **128**, 1761

[2] J. Agassi, This *Journal*, 1963, **14**, 26

[3] In the semi-classical theory the field is not quantised, only the interaction with the detector, see, for example, M. Born and E. Wolf, *Principles of Optics*, Oxford, 1959, chapter x.

[4] E. C. G. Sudarshan, *Phys. Rev. Lett.*, 1963, **10**, 277

[5] S. I. Vavilov, *Uspekhi fiz. nauk*, 1936, **16**, 892

Dept. of Physics,
Imperial College,
London

DISCUSSIONS

TWO FORMS OF THE PREDICTION PARADOX

IN an article by one of the authors in *Mind*[1] it was argued that the prediction paradox, in the form there given, could not be resolved within the framework of classical two-valued logic because, it was claimed, one was able from true premisses to draw a false conclusion. It is shown below that this is not so and only appeared to be so because of an insufficiently sophisticated use of logic.

This form of the paradox can be stated as follows: On Sunday a captain has decided to give his men a route-march on the following Wednesday, but merely tells them that on one of the six following days he will give them a *surprise* route-march. He defines this as a march for which the day of occurrence cannot validly be deduced beforehand. The argument now goes that the march could not occur on the last day, Saturday, because on Saturday morning it would be deducible that it would occur on that, the then only possible, day and therefore would not be a surprise. Nor could it occur on Friday, since on Friday morning it would be known that Saturday was ruled out, that therefore it would occur on Friday and so would not be a surprise. Similarly it can be shown, in succession, that it could not occur on Thursday, Wednesday, Tuesday and Monday. Nevertheless, it *does* occur on Wednesday and, apparently, this was not deducible. So, apparently, the captain's statement was true but a false conclusion could be drawn from it—namely, that the march would not take place at all.

But in fact the captain's statement was *not* true. It was a conjunction of two parts: the first asserted that there would be a march—which was indeed true; the second asserted that the day on which it occurred was not deducible beforehand. This is not true, for the argument given above has simply not been pushed far enough. It was there shown that the march could not take place on Monday, Tuesday, Thursday, Friday or Saturday. On the other hand the premisses stated that the march would occur on one day of the week. It follows immediately that the march would occur on Wednesday, so that the day of occurrence *was* deducible.

This reasoning has an appearance of legerdemain about it, but this is usually so in applications of the valid logical principle that from a contradiction one can deduce any proposition whatsoever. A well-known example is Bertrand Russell's deduction from an arithmetical falsehood that he is the Pope.

The need for the use of this principle disappears if we consider the one-day form of the paradox, as Quine[2] does. In this case the captain states there will be a march today and it is not deducible that this is so. This is immediately seen to be a false statement because from a proposition A one can certainly deduce A.

---

[1] Meltzer, B., 'The Third Possibility', *Mind*, **73**, 1964, pp. 430-433

[2] Quine, W. V., 'On a So-called Paradox', *Mind*, 1953, **62**, pp. 65-67. Quine is concerned with a problem different from ours, since it is for him a matter of what the soldiers *know*—a matter that goes outside pure logic. As he shows, the form of the problem that he considers is not paradoxical.

It is interesting to consider another form of the paradox, in which one changes the definition of a surprise march. The ordinary meaning of 'surprise' involves the notion of probability, and if this ordinary meaning is used, the resolution of the paradox becomes trivial. The captain used a non-probabilistic interpretation of surprise, as described above, but there is another non-probabilistic interpretation that he could have used, perhaps somewhat more artificially. We shall now say that a march will be termed a surprise if it occurs on a day on which it is deducible beforehand that it will not occur! We may now again argue that the march could not occur on Saturday; because if it had not occurred by Friday night it would *not* be deducible that it would not occur the next day, so that the Saturday march would not be a surprise. And, as before, one can argue back day by day and show that all six days are ruled out. So in this case the captain's statement is apparently true, since the march *does* take place on Wednesday and it has been deduced beforehand that it will not take place on Wednesday (among other days). Nevertheless from this apparently true statement an apparently false conclusion has been drawn, namely, that the march does not take place on Wednesday.

The authors do not agree on the resolution of this paradox. One would suggest that a three-valued logic is required, on the lines that he discussed in *Mind*.[1] But the other takes a view regarding this logical antinomy, and most others, which can be best presented by considering the one-day form of the problem:

The captain makes a statement S which is of the form: 'A, and not-A is deducible from the whole of this statement.' Here A is meaningful and true by hypothesis, and, if S were meaningful, then an argument similar to that of the last paragraph but one shows that it would be the conjunction of two true statements and would therefore be true, and could not lead to contradictions. Since, however, it does lead to contradictions, it follows that it is meaningless. But the conjunction of two meaningful statements is meaningful (this may be taken as an axiom of semantics), and A is meaningful; therefore both S and the second clause of S are meaningless.

B. Meltzer and I. J. Good

The Logic and Psychology of Science:

Reply to Professor Kapp

In his review of my book, *The Origins of Science* (this *Journal*, 1965, **15**, 333), R. O. Kapp ascribes to me a view which is exactly the opposite of what I have said. On page 135, I have written of Euclid that, 'He was the first to overcome the horror of the infinite and to show that it could be brought under some sort of control: we speak today of the "countable" or "denumerable" kind of infinity.' Since the Greeks had been occupied with the problem of infinity ever since they began to speculate about nature, it is absurd to suggest that 'Euclid was the first to be overcome by the horror of the infinite', as is done in Kapp's review (p. 340).

---

[1] Meltzer, op. cit.

There is also another misinterpretation of my text. I say about the Mind-Body Problem, ' There is no problem in *the ordinary sense* of the word,' not—as Kapp accuses me of saying (p. 338)—that there is no problem at all. On the contrary, I assert that it is an emotional (or psychological) problem and, indeed, that it is the basic ' problem ' underlying Western philosophy.

Finally, I want to point out that the material contained in this book was presented in public lectures as early as 1957. This is stated clearly in the preface. Thus, there was no need for me to mention Koestler's *Sleepwalkers*, a book which was published in 1959.

E. H. Hutten

REVIEWS

WHAT IS HISTORY?

PROFESSOR GALLIE's book[1] consists of two parts: the first is concerned with an account of the nature of historical understanding, the second with a discussion of the importance of historical understanding to the practice of politics, science, and especially philosophy. While the two parts are interrelated, the first is of greater relevance to the problems usually regarded as belonging to the philosophy of history, and it is at times very stimulating and provocative. I shall, therefore, confine my remarks to the first part, a shorter version of which appeared in *History and Theory*, 1963, **3**. Gallie claims that what is new in his account of historical understanding is the emphasis he has placed upon the idea of a narrative. There are, as Gallie acknowledges, anticipations of his approach to the philosophy of history in the writings of Popper and Berlin; a more comprehensive history of the idea that the writing and comprehension of a history resembles significantly the telling and following of a story would include mention of the writings of Carl Becker and more recently of Arthur C. Danto (' Narrative Sentences ', *History and Theory*, 1962, **2**).

With the story rather than the sciences as his guide, Gallie sets out to correct the mistakes and faulty emphases he finds in the writings of those philosophers who believe that an adequate explanation must conform to the ' deductivist ' or covering law model. His emphasis upon the pragmatic aspects of explanation at the expense of formal considerations leads him to stress, unduly in my opinion, the differences between explanation and prediction and to mix and perhaps to confuse the questions that usually occur in discussions of ' determinism ' and ' contingency ' with those that usually arise in discussions of the character and structure of historical explanation. It may be that a critical discussion of the ontology of covering law theorists would be fruitful (certainly attention should be paid to their reluctance to take seriously attempts to distinguish between an event in nature and an event in history); but the way in which Gallie makes a case for there being ' unpredictable events ' is more disturbing than fruitful. It is not always clear whether he means to say that there are some events which are unpredictable in principle or that there are some events which, given our lack of appropriate information, are as a matter of fact unpredictable by us. At times he seems to be saying both. The first claim would certainly be upsetting to the covering law theorist or to anyone who seeks a rational account of the world about us. The second claim need not bother anyone; the covering law theorist has always recognised that in many situations our predictive powers may be quite weak. We may lack sufficient knowledge of the relevant circumstances, and we may lack a general law or laws that would 'cover' the phenomenon before us. The covering law theorist is only claiming that the form of prediction and explanation is the same. His claim is not that prediction is as easy in all cases as explanation may be; nor does he claim that anything I can explain now I could have predicted before its occurrence regardless of my state of knowledge at the time prior to its occurrence. Rather the claim is that in cases where I have successfully explained a given phenomenon, 'if I had known then what I know now' I could have successfully predicted its occurrence.

---

[1] *Philosophy and the Historical Understanding*, by W. B. Gallie, Chatto and Windus, London, 1964, pp. 236, 30s.

Because of the importance of contingency, coincidence, and accident in the historical process the historian is, according to Gallie, more interested in presenting his readers with a plausible narrative of events than with either the explanation or the prediction of events. To be sure, the historian employs generalisations, uses general terms, and has an eye for the recurrent or typical, but these are the recessive and uncharacteristic features of his activities. Explanation in history functions like explanation in a story. It seldom occurs, and when the author stops to explain he does so only to clear up certain difficulties or complexities that may baffle the reader. Whether one is following a game, a story, or a history, explanation figures as an intrusion and not as an integral part of the narrative. The really important thing about a game, Gallie tells us, is not its rules but the spirit and skill with which the players deal with one another and with various contingencies; the really important thing about a story or a history is not the explanations the author may offer but whether he has provided a followable narration of events moving toward an 'open' end (an end which is, however, made to seem plausible and convincing as the story or history unfolds).

The thesis that in some vital respects stories are like games and histories are like stories is illuminating, although I should prefer to say that some stories are like some games and some histories are like some stories. In any event, Gallie's treatment of this game-story-history theme strikes me as curiously incomplete; it seems to presuppose an especially sharp distinction between telling and explaining which simply is not borne out by the stories and histories I read. Perhaps Gallie reads different authors. Perhaps there are different kinds of stories and of histories. There are certainly different techniques. In the novel, for example, it is well known that some authors speak through an omniscient narrator, others use an obtuse narrator who almost understands what is happening, and others use stream-of-consciousness techniques in which the characters often can only offer clues as to why they are behaving as they are. My impression is that in all of these cases, and others as well, a great deal of explaining is being done in a variety of ways. When John O'Hara goes into exhaustive detail about the house, clothing, and transportation facilities owned by one of his characters it is hard to say where the describing and the telling leave off and where the explaining begins. Mightn't the same be true of the historian and his narratives? Isn't there a difference between a *pat* explanation and one which is essential to our comprehension not only of why certain people behave as they do but even to our understanding of what it is that they are doing?

Regardless of whether histories are like stories in their structure and dynamics it seems obvious that we ask of histories questions concerning their truth in a way in which we do not ask such questions of works of the imagination. Where their ' explanations ' are concerned a novelist and a historian might be equally skilful in explaining certain events, but the historian's reputation depends upon the truth of his explanations in a way that the novelist's does not. D. H. Lawrence's reputation as a storyteller is not going to be seriously affected by what posterity thinks of his thesis that industrialism corrupts both man's art and his sexuality, although this thesis is an integral part of many of his stories and essential to our understanding of their movement. On the other hand, Charles A. Beard's reputation as a historian seems largely dependent upon what scholars think of his ' economic interpretations ' of the American Constitution and the American Civil War. If, as Gallie claims, ' explanations are not the main goal and the main hall-marks of achievement in history ' (p. 123), mightn't they be *one* of the main goals and main hall-marks of achievement in history? Mustn't they be one of the main goals, in view of the earnest debates among historians (in Europe and America more than in England) concerning the explanatory value of various economic, psychological, and sociological ' interpretations ' of history and in view of the fact that a historian's reputation will often rise or fall according to his ability to explain or to reinterpret a portion of the past in the light of a new hypothesis, either one of his own devising or one borrowed from the social sciences? The history of history is often written in terms of ' schools ' which may owe their existence largely to their commitment to a given historical interpretation or explanation. Gallie might reply that this is an unsubtle way of writing the history of history, but it is after all the historian not the philosopher who decides how history is to be written.

The history of history shows something else which is, or should be, unsettling to Gallie's argument that the narration of a story as he has analysed it provides an adequate model for understanding what the historian is doing. The history of history reveals a distinction that was once current among American historians between ' literary ' (narrative) historians and ' scientific ' historians. I think that some such distinction is still worth making. In America at least there are some productive professional historians who do not write narrative history and who look down upon those historians who still profess to be story-tellers. Is this division of no consequence? Have the scientific historians merely been led astray by the social sciences, or is it the case, as they have contended, that history is becoming more complex and that the story-tellers are on the way out?

One of the peculiarities of historical explanation is that in some respect almost everyone seems dissatisfied with it. On the one hand, there are those covering law theorists who complain that the historian does too little, that his explanations are incomplete sketches of the lines that a more comprehensive explanation might follow. On the other hand, there are those who like Gallie appear to feel either that historians sometimes attempt too much or else that by the use of misleading expressions they delude some philosophers into believing that the historian is aspiring to be a junior-grade scientist. Gallie's dissatisfaction with some of the historian's linguistic behaviour is evident in that passage (pp. 110–111) where he argues that in historical writings ' what appear to be explanatory sentences—or what might be taken to be such by overzealous logicians—can often perfectly well be replaced by a number of narrative sentences which no one would dream of regarding as explanations.' Gallie briefly gives two examples of historical statements in which, he thinks, such replacements or substitutions would be feasible. In his first example, he proposes the elimination of a ' therefore ' and in the second of a ' because '. I shall consider only his first effort in what seems to be the simpler case, proceeding on the assumption that if his strategy doesn't work here it probably wouldn't work in the more complicated case either.

Gallie invites us to consider this statement: ' The news had now reached the President who therefore gave the following orders.' He offers this analysis: 'Is it not clear that . . . the "therefore" means little if anything more than " thereupon "?—i.e. it points to an intelligible or followable next step, but makes no claim that the main clause is or contains a sufficient condition of the subordinate clause which follows. Nor, I think, is there any presumption that the main clause expresses a necessary condition of what follows. Conceivably the President would have given the same order if slightly different news, or even if no news, had come through.' My suspicion is that when a conscientious historian says ' therefore ' he means ' therefore ' and not ' thereupon ' and that Gallie is probably guilty of special pleading on behalf of ' narrative sentences '. Of course, the news that reached the President could not by itself be a sufficient condition of his delivering any orders. Also, it is ' conceivable ' that he would have given the same orders if he had received slightly different news or if he had received no news at all. Suppose, for example, the news he received was that one of his reconnaissance planes had been shot down instead of two; or suppose that he didn't receive any news from the reconnaissance planes, that they simply disappeared, and that he had decided to issue orders for an attack regardless of whether any of the planes returned. But it is also conceivable that he might have said (and meant), ' I won't attack *unless* the reconnaissance planes report A, B, and C.' And suppose that the reconnaissance planes do report A, B, and C. Doesn't this make the fact that certain news has reached the President a necessary condition for his issuing the orders to attack? Isn't it highly probable that this is the sort of thing a historian would have in mind when he writes that the President, therefore, gave such and such orders? I doubt that the careful historian's use of words such as ' therefore ' and ' because ' is as haphazard as Gallie implies.[1]

If, following Gallie (and Danto), we are to take a new look at the logic of narrative sentences and their role in historical writing, we must take great pains to avoid or at least to suspend the belief that narrative sentences can do, and do better, the jobs that explanatory sentences do. It may, of course, still turn out that despite the failure of Gallie's strategy in the above example workable narrative substitutions might be made in *some* cases. Maybe some ' therefores ' really are covert ' thereupons ', etc. But this in turn might seem to be a significant discovery only because Gallie has overlooked from the outset the ties and resemblances between explanation and description and has assumed too sharp a distinction between them.

BURLEIGH TAYLOR WILKINS

The Anatomy of Inquiry. By Israel Scheffler.
Routledge & Kegan Paul, London, 1964. Pp. xii + 332. 45s.

THIS is the second in the series of Borzoi Books in the Philosophy of Science. In the Introduction, Professor Scheffler explains that it is directed towards a theory of scientific structure, that is, towards a systematic and philosophically intelligible account of such structural terms as ' explanation ', ' significance ', and ' confirmation '. In Part I he discusses the deductive pattern of explanation (that is, what is sometimes called ' the covering-law model '), considers whether prediction and explanation are structurally alike, and examines teleological explanation as a possible rival to the deductive pattern. In Part II he examines various attempts to frame a criterion of empirical significance, compares with these Popper's use of falsifiability as a criterion for the demarcation of science, and discusses pragmatism and several varieties of fictionalism as ways of coping with the fact that science seems to require, in its systematic theories, non-observational terms which lack 'intuitive clarity'. In Part III, after sketching 'the challenge to justify induction, arising out of Hume's analysis', he discusses at considerable length Hempel's proposed definition of confirmation and the associated paradoxes of confirmation, and concludes with a review of Goodman's theory of projectible hypotheses and some final comments on the justification of induction.

---

[1] In an earlier paper Gallie argued that historians explain events by referring to one or more of their temporally prior necessary conditions and that in this respect historical explanations are like explanations in the genetic sciences (' Explanation in History and the Genetic Sciences,' *Mind*, 1955, **64**). While to my mind historians are interested in both the necessary and sufficient conditions of events, Gallie's earlier paper had the advantage that it did not minimise the importance of explanation in historical writing.

Since he is concerned with questions of structure which apply equally to different branches of science, Scheffler does not go into the problems of any specific science, and the examples he uses are simple, even trivial: as he admits, they belong rather to everyday life or to simple investigations than to theoretical science. In fact, his interest lies in the logical tangles and in the problems of formulating criteria rather than in finding out how these structural concepts are used in actual scientific work or in resolving any difficulties that arise there.

Scheffler's method consists largely in reporting, developing, and commenting on discussions that have been published over the last twenty years or so. He is indebted particularly to Hempel and Goodman, and both on explanation and on confirmation the general line of his argument is that Hempel's views can be defended, various criticisms of them being rebutted, or shown to apply to different questions, or accommodated without any radical change in Hempel's position. Much of what Scheffler says is sound, but his style is heavy, his method of discussion laborious. For these reasons, and because of his meticulous concern with the details of logical formulation, this book will not serve to introduce beginners to the philosophy of science; on the other hand, it contains too little that is new or independent to be of great interest to those who are already familiar with the subject. Nevertheless, the advanced student will find it useful as a collection of, on the whole, careful and balanced discussions.

Accepting the limitations of Scheffler's aim and method, and admitting that there are in this field acute problems which no one has so far been able to solve, this book is still less satisfactory than it might have been. It suffers from two outstanding weaknesses: a tendency to sidestep important issues and a failure to bring different considerations together into a unified theory.

For example, Scheffler traces various attempts to give a criterion for the empirical significance of statements in terms of verifiability and falsifiability, direct or indirect, and shows that they fail, on formal grounds, to give the discriminations we require. But he does not see that what this reveals is that we can and do make discriminations between empirically significant and non-significant statements on principles that are not tied to checking procedures at all, and that it would be better to go back to the older notion of accounting for empirical significance by describing ways in which terms and statements may be constructed out of observational elements, and of excluding as non-significant items for which no satisfactory construction can be offered. It is, after all, by an implicit appeal to this principle that we know that statements in which the term ' glubbified ' occurs essentially are non-significant, a fact on which Scheffler repeatedly relies in formulating counter-examples to various proposed tests of significance. Scheffler comes close to this more fruitful approach when he considers translatability into an empirical language as a criterion of significance, but he dismisses this suggestion for what appear to be extremely shallow reasons (p. 169). If this approach were adopted, there would no longer be a problem about dispositional terms or the subjunctive conditionals they involve. For these conditionals can be understood as telescoped arguments, some of whose premises may be universal propositions:[1] and there is no difficulty about the significance of universal propositions, provided that their terms have been given empirical significance, and no great puzzle about the use of them in arguments.

Again, Scheffler assumes that some terms which are important ingredients in systematic science lack ' intuitive clarity ', that is unquestionable literal meaning, and goes on to examine pragmatism and fictionalism as alternative ways of dealing with this fact. However, since they are distinguished on the ground that pragmatism grants whereas fictionalism denies significance to items that figure in an effective systematic account but lack intuitive clarity, the disagreement between these two attitudes is minor, perhaps merely verbal, whereas there is a major issue between both of these on the one hand and realism about theoretical entities on the other. But Scheffler simply does not discuss this major issue : yet if what I described above as the more fruitful approach to the question of significance were adopted, it might be possible to defend the literal significance of at least many theoretical terms and a realist view of the corresponding entities.[2]

However, if there are in scientific theories any important items, a realist view of which cannot be defended, then some other account of these terms is needed, and Scheffler's discussion of the Ramseyan form of eliminative fictionalism is one of the most interesting sections of this book. The Ramseyan method consists essentially in replacing all non-empirical constant terms, including predicates, by existentially quantified variables. Scheffler's conclusion is that though the Ramseyan substitute for an ordinary theory has the same observational consequences as that theory itself, the original theory has ' greater inductive utility '. This conclusion is implausible, since we can always go by existential instantiation from the Ramseyan substitutes to something that is *formally* identical with the original theory : but in any case Scheffler's argument, though complex, is not and cannot be convincing, because he has not yet formulated any account of inductive confirmation, and indeed he never reaches any clear and unified account of this.

---

[1] See, for example, my article ' Counterfactuals and Causal Laws ' in *Analytical Philosophy*, Oxford, 1962.

[2] Cf. J. J. C. Smart, *Philosophy and Scientific Realism*, London, 1963, Chapter ii

In Part III, Scheffler admits that there is some force in each of the various treatments of the paradoxes of confirmation—Hempel's view that any observation that satisfies a hypothesis confirms it, and that any appearance of paradoxicality in this is an illusion arising from the illegitimate use of further information, Watkins's view that confirmation results only from *tests* of an hypothesis, and the view of Hosiasson-Lindenbaum, Pears, and others that the degrees of confirmation of an hypothesis by observations of various sorts depend upon the relative size of the classes involved. But he fails to bring these different suggestions within any general theory. He suggests that Hempel's theory is satisfactory as an account of what a positive instance is, whereas Watkins's is relevant to the question of what adds strength to a hypothesis at a specified time for a given person. Scheffler criticises the class-size theory unfairly (p. 284), for example on the ground that if invertebrates are not fewer than things not lacking kidneys, the family cat should, on this theory, be a positive instance of the hypothesis 'All invertebrates lack kidneys': but whether we call the cat a positive instance of this hypothesis or not—and on Hempel's view as Scheffler defends it, we must—there is in fact nothing paradoxical in taking an observation of the cat as confirming this hypothesis, and confirming it better than the observation of one jellyfish. But what Scheffler misses is that if the notion of testing a hypothesis is clarified, these three views can all be explained as applications of a single principle—the principle that an observation confirms a hypothesis if and only if, or in so far as, the observation-report is made more probable by the adding of the hypothesis to the background knowledge. Applications of this principle to different contexts determined by different assumptions about the available background knowledge will yield the three different treatments each of which Scheffler admits to have some force.[1] This method can even be extended to cover Scheffler's claim (pp. 284-285) that the paradoxes 'involve intuitive inequalities (among reports) *which themselves vary* with respect to logically-equivalent versions of " the same hypothesis ".' If we are inclined to take a black raven as confirming 'all ravens are black' but not as confirming 'All non-black are non-ravens', this may be because these different formulations of the hypothesis, though equivalent in what they assert, may not be equivalent in the background knowledge they tend to presuppose, and there is nothing paradoxical in the same hypothesis being confirmed in relation to one background, but not in relation to a different background, by the same observation. What we can do, therefore, is not merely to clarify and distinguish several different senses of 'confirmation': we can develop a unified general account of confirmation which has different consequences in different contexts. Moreover such a general description of confirmation is what we need if this is either to replace attempts to justify induction, or to serve as the first stage in a new attempt.

[1] See my article, 'The Paradox of Confirmation', in this *Journal*, 1963, **13**, 265-277.

A similar lack of unity affects the general structure of this book. Although one topic runs on into another, Scheffler does not build any systematic theory himself, he does not use in one section the conclusions reached or the concepts developed in another.

Some specific weaknesses should also be noted. Scheffler's sketch of the problem of induction (pp. 227-236) is far too hasty: he does not examine the way in which a principle of uniformity actually enters into inductive reasoning—namely as an assumption in eliminative induction—and he accepts far too easily both the common view that any justification of induction must be circular and the quite unwarranted charge that every search for a justification is a demand for a deductive rationale. Correspondingly, he finally accepts Goodman's purely conventionalist view that 'Predictions are justified if they conform to valid canons of induction ; and the canons are valid if they accurately codify accepted inductive practice' (quoted on p. 317); but whatever we ought to say about induction it is certain that *this* won't do.

Again Scheffler brushes aside far too casually such criticisms as Dray's of the application to history of Hempel's theory of explanation (pp. 76-80). We can indeed distinguish 'explanation sketches' (gappy or incomplete explanations) from complete deductive explanations. But it is important to note that such sketches can be supported and even established without being completed, and that we can even confirm the corresponding incomplete laws without completing them. Gappy explanations and gappy laws are often all that we can discover, and they can be and are used both in science and in its applications : they do not serve *merely* as preliminaries to or summaries of complete deductive explanations. An associated point is that while Scheffler detects a minor divergence between explanation and prediction (pp. 43-46) he misses a much greater divergence. We can frequently give an explanation sketch, stating *and justifying* the claim that one event occurred because of another, where the actual occurrence of the later event is a vital part of the confirmation of this explanation sketch and of the incomplete law that is implicit in it, and where we therefore could not have predicted this event from the *rest* of the available information.[1]

J. L. Mackie

Foresight and Understanding. By Stephen Toulmin.
Hutchinson, London, 1961. Pp. 115. 18s.

In the first chapter of this very well written little book Stephen Toulmin argues against various forms of the thesis that the aim of science is prediction. He concludes that all forms of the thesis fail, since the aim of science is explanation and the prediction theories can give no adequate account of this. However, the vagueness of Toulmin's own account of explanation squares ill with the rigour of his criticisms of the prediction thesis. To explain is to 'make sense' of the observed facts, to provide principles in terms of which they 'hang together', to make what before appeared 'arbitrary' appear 'natural and rational'. This is done by means of 'Ideals of Natural Order', and in the following two chapters two pieces of history, that of theories of motion and that of theories of matter, are presented in the light of the various 'Ideals of Natural Order' which inspired them. These pieces of history are well done, bringing out the fundamental assumptions of the scientists under the title of their Ideals. How Toulmin intends them to further the understanding of the nature of explanation is not clear.

---

[1] This point is made by Michael Scriven in a review of Nagel's *The Structure of Science*, in *Review of Metaphysics*, 1964, 17, 414. But Scriven wrongly infers that this is 'disastrous for the covering-law theory of explanation'. It is disastrous only for the assimilation of explanation to prediction: the covering-law model can be retained if it is admitted that a covering-law may be used in an incomplete form.

The other task Toulmin sets himself is to account for the different ways in which theories excel. He presents a reformulation: Theories excel which outlive other theories in the battle for survival; what gives theories survival value? Now just as there can be no criteria for a well adapted species, since species adapted in one environment may fail to survive in another, so there can be no criteria for excellence of a theory, for a theory which explains the facts as we know them at present may be unable to adapt itself to new discoveries. Theories of confirmation are doomed to failure. But their opponents fare no better. The view that confirmation must fail because any set of observations confirms an infinite number of mutually incompatible hypotheses is irrelevant to scientific work. Far from the scientist having an infinite number of hypotheses to choose between, 'it may be a stroke of genius for him to imagine even a single one'.

A serious error of fact mars the chapter on 'Forms and styles of theory'. Toulmin quotes a passage from Kant's *Prolegomena* concerning Newton's inverse square law of gravitation:

> . . . we find a physical law of reciprocal attraction applicable to all material nature, the rule of which is that it decreases inversely as the square of the distance from each attracting point—that is, as the spherical surfaces increase over which this force spreads—which law seems to be necessarily inherent in the very nature of things. . . .

Toulmin takes Kant to be arguing from the premiss that a given mass provides a fixed amount of force which when spread over a surface of larger area gives a smaller force at any point of that surface. Toulmin draws an analogy with the proof that intensity of illumination varies inversely as the square of the distance from the point source. Whilst this latter proof, according to Toulmin, is valid, Kant's proof is not: 'Kant's proof acquired neither a pragmatic value nor a theoretical basis.' It requires us to believe 'that gravitational force spreads out from the sources of attraction like an indestructible fluid'. Its validity depends upon the assumption 'that gravitational force is propagated and conserved, in the way that matter and energy can be'.

In fact the premiss which Toulmin takes to be Kant's, and to which he denies pragmatic value, has a fundamental place in the theory of gravitation; it is Gauss's Theorem:

> The outward flux of the force of attraction over any closed surface in a gravitational field of force is equal to -4π times the mass enclosed by the surface. (A. S. Ramsey, *Theory of Newtonian Attraction.*)

Few theorems have had more pragmatic value than this, and if it has no theoretical basis it certainly provides one. The theorem gives the inverse square law immediately and is itself a consequence of that law together with the assumptions that provide for the applicability of the integral calculus in the theory of gravitation. Of course Kant's beliefs that the theorem is *a priori* valid, and that it has greater self evidence than Newton's law itself, are other questions. What mistake leads Toulmin to reject Kant's premiss as a mistake? Since the theorem concerns *any closed surface*, however distant from the massive bodies it contains, it does assert what Toulmin expresses by saying that 'gravitational force spreads out from the sources of attraction like an indestructible fluid'. So Toulmin's argument goes like this: Kant's premiss implies the indestructibility of gravitational force. This in turn implies the conservation of gravitational force. But though energy is conserved, force is not, therefore Kant's premiss is false. In fact, since Kant's premiss, Gauss's theorem, is true, and since force is not conserved, then there is nothing left but to reject Toulmin's belief that obedience to Gauss's theorem entails the conservation of force. There are independent reasons against this belief, for the indestructibility of Gauss's theorem is indestructibility throughout space, whilst conservation is throughout time.

JOHN WATLING

The Development of Logic. By William and Martha Kneale.
Clarendon Press, Oxford, 1962. Pp. viii+761. 75*s.*

NOT only is *The Development of Logic* most valuable as a work of reference, but it is also a delight to read because it sets a rare standard of excellence in so many ways. In spite of being concerned with academic philosophy, it is entirely free from the prevalent weaknesses of much philosophical writing: aridity and excessive technical elaboration in mathematical logic, triviality and looseness of thought in philosophical logic, and archness, ostentation, and slickness in philosophical discourse in general. The book is serious, and the argument in it is never idle; but technical means are used with restraint, and the superior intelligibility of ordinary language is not sacrificed unnecessarily. Indeed, natural languages are seen as fundamental instruments of thought, which have many functions besides that of serving as primitive systems of logical symbolism. The Kneales' literary style is wholly admirable; and though their writing is quiet by comparison with that of such authors as

Collingwood and Whitehead, it is always vigorous, and its easy flow does not cease to charm the eye and ear even after many re-readings.

The authors' choice of theme and subject-matter is a personal one, but in retrospect it has an inevitability that makes it seem absolutely appropriate. The theme is stated in the title of the book; and we are in fact shown how logic came into being in ancient times, persisted with no great change for many centuries, and then surged forward when logicians discovered the relevance to their studies of the concepts and methods of mathematics. By studying historically the two great ages of logical inquiry—the one dominated by Aristotle and the other by those logicians of whom Frege is perhaps the greatest representative—we are enabled to appreciate the rich complexity of logical theory, and to disentangle its ramifications.

It goes almost without saying that Aristotle was by no means the only logician of his time, although he was by far the most important. Inadequate documentation makes the detailed history of ancient logic partly a matter of speculation, but some of the more significant developments can be traced with reasonable confidence. Martha Kneale's chapters on the logic of the Greeks contain many interesting discussions of this kind, which bring out clearly the common purpose which the logicians of antiquity shared with modern symbolic logicians, and also the limitations to which early logic was subject by reason of the lack of an adequate technique for the handling of formal relationships.

It is especially significant that although *The Development of Logic* is the work of two philosophers, neither of them professionally concerned with mathematics, nearly half of its 750 pages are devoted to 'modern' logic, that is to say logic into which mathematical ideas have been infused. William Kneale, who is responsible for this part of the book, does not look upon mathematical logic as a narrow study, of interest only to mathematicians, and capable of dealing only with mathematical subject-matter. For him, mathematical logic is simply logic, in the form that is appropriate to the present age. 'Logic is concerned with the principles of valid inference'; and since mathematical concepts and principles are an inseparable constituent of the fabric of modern thought, any logic that fails to treat them seriously is thereby condemned from the outset to irrelevancy. To have conveyed this conviction is perhaps the Kneales' main achievement. Their scholarship is impeccable, and their narrative is lively and beautifully phrased; but what matters most of all is that they establish authoritatively but without any fuss, and more by their own attitude than by explicit declaration, the essential truths that logic has at length come of age, that it is an equal partner with mathematics in disciplining thought and providing it with the formal equipment necessary for the discharge of its cognitive function, and that the distinction between two modes of logic, the one philosophical and the other mathematical, must now be deemed obsolete.

G. T. KNEEBONE

The Principles of Science, A Treatise on Logic and Scientific Method. By W. Stanley Jevons.

Dover Publications, 1958. Pp. liii + 786. $2.98

This is a paper-backed reprint of the well-known work of Jevons which appeared first in 1874. The setting of the pages, which reproduces exactly the lay-out of the second edition (stereotyped in 1877), is presumably the result of a photographic process. In a new introduction Professor Ernest Nagel gives a sketch of Jevons's life and explains briefly his contributions to (i) the class calculus, (ii) the theory of ampliative induction as an application of hypothetico-deductive method, and (iii) the theory of probability. He makes it clear, of course, that in these three fields Jevons elaborated the work of Boole, Whewell and Laplace, and he points out that Jevons's development of what he inherited was not in all respects perfect. But he rightly stresses the point that the attempt to bring together the three notions mentioned above was an important step in the history of the philosophy of science. Perhaps no one since Leibniz had thought of all these topics together. We now know that a formal logic without quantifiers is inadequate for the purposes of science, and we are properly sceptical about the thesis that ampliative induction can be justified by a probabilistic argument from the principle of indifference; but it is evident to us that a really comprehensive book about the principles of science would have to include all the topics Jevons discussed. On the other hand, as Professor Nagel remarks, it is rather curious that Jevons did not take note of developments in the philosophy of geometry about which he might have learnt from Clifford, and that, unlike John Stuart Mill, he did not consider the peculiar problems of the social sciences in which he was himself a distinguished practitioner.

In their advertising material the publishers draw special attention to Jevons's account of the logical machine which he exhibited to the Royal Society in 1870. It is pleasant to record that the machine is now safely housed in the Museum of the History of Science at Oxford, intact except for the catgut by which its moving parts were originally connected. If computers acquire from their creators a disposition to ancestor worship, we may some day see them queuing in Broad Street to pay their respects to this simple forefather of their race.

William Kneale

Historical Studies in the Language of Chemistry. By M. P. Crosland.

Heinemann, London, 1962. Pp. xvii + 406. 50s.

Dr Crosland has very successfully presented in this book a history of chemistry from earliest times up to 1892 by tracing the evolution of the terminology and language of the science. Such a book could easily have become a heavy catalogue of multitudinous facts, but the author has avoided this pitfall. His book is readable, extremely interesting to the chemist, and very learned. It is well documented, and is supplied with an extensive bibliography and good index.

Different readers will turn and return to Dr Crosland's book for different reasons. I should like to mention some of the themes which I found particularly interesting. The Alexandrian chemistry of the metals (p. 80), the 'New London Pharmacopoeia' of 1745 with most interesting references to red precipitate (p. 117), the influential work of the somewhat neglected genius J. B. Richter (p. 209), Dalton's symbols for the elements (p. 259), and the history of that most famous of all chemical Congresses held at Karlsruhe in 1860 (p. 344) are all presented with distinction, and there will be few students of the history of chemistry who will not find something new in Dr Crosland's work.

The author has inevitably been selective, and there are some parts of his book which could mislead the more inexperienced reader. He is, I think, too brief on Lavoisier and Lavoisier's *Traité* (p. 190), and perhaps somewhat unsympathetic to the phlogistonists and the influence of their great theory on the terminology (p. 170). The last main section of the book (p. 285) is mainly concerned with the language of organic chemistry in the nineteenth century, and I think here Dr Crosland has been to some extent defeated by the intricate complexity of the story. His very condensed account will help the advanced student, but is more likely to baffle the beginner.

There is no great philosophical interest in the method of treatment, but it is a book from which philosophers of chemistry could draw valuable lessons. There are a few misprints and other unimportant errors. There are interesting and good plates and diagrams.

JOHN BRADLEY

ABSTRACTS

Philosophy of Science, 1964, 31, No. 1

J. Losee, 'The Use of Philosophical Arguments in Quantum Physics'

Two types of philosophical arguments are employed by the defenders and critics of the Copenhagen Interpretation. One type of argument is a confrontation of an opponent's interpretation with criteria of demarcation and criteria of acceptability. The purpose of such arguments is either to exclude an opponent's interpretation from the range of permissible discourse in quantum physics, or to establish the inadequacy of an opponent's interpretation. A second type of argument is a justification of the value, or utility, of the criteria selected.

A. Edel, 'The Concept of the Unconscious: Some Analytic Preliminaries'

An attempt is made to disentangle issues in controversies about the unconscious. Eleven questions are distinguished and discussed. Logical, linguistic, methodological, metaphysical, empirical, and pragmatic components are set apart. It is found that there are no logical barriers to a construct of the unconscious, that it is linguistically feasible, need violate no methodological concepts, nor foreclose a metaphysical issue, nor have deleterious pragmatic effect. Feasibility and desirability of the construct in all these respects will depend on the character and extent of the scientific findings. Desirable types of models in the explanation of conduct are discussed, as is also the relevance of recent studies of consciousness.

P. R. Wilson, 'On the Argument by Analogy'

Conditions are stated under which 'argument by analogy' is consistent with the principle of inverse probability. It is contended that argument by analogy, in conjunction with a crucial test, has a legitimate place in scientific logic. As an example the astrophysical problem of solar granulation is discussed in detail and other examples are mentioned more briefly.

J. Rothstein, 'Thermodynamics and Some Undecidable Physical Questions'

A number of questions, usually considered philosophical rather than scientific, can be reformulated to apply to a world of automata or 'well-informed heat engines'. In some cases they admit of physical answers, but in many cases obtaining answers entails violation of the second law of thermodynamics. This is demonstrated explicitly for the problem of determinism and free will, for the discovery of the origin or ultimate fate of the universe, or for the discovery of causes or purposes in nature.

A. Pap, 'Theory of Definition'

Philosophy of Science, 1964, 31, No. 2

W. Todd, 'Counterfactual Conditionals and the Presuppositions of Induction'

Professor Goodman was correct in thinking that there is a problem concerning counterfactual conditionals, but it is somewhat different from the problem he thought it to be, and is one that is even more basic. This problem is distinct from Hume's 'problem' of induction, and that additional assumptions have to be made for counterfactual induction beyond those required for other kinds of induction.

D. Shapere, 'The Casual Efficacy of Space'

Through an analysis of conditions under which the question of spatial anisotropy can be raised, the present paper brings out intimate conceptual relationships between the scientific concept of space and the concepts of entities, behaviour, and explanation specified by scientific theories. Thus scientific departures from ordinary usage (or from usage in other scientific theories) of the term 'space' entail corresponding shifts in the use of other terms not generally seen to be connected. As à case study of the relations between these terms, and of the refusal to allow the possibility of spatial anisotropy, Newtonian mechanics is examined.

F. I. Dretske, 'Particular Reidentification'

A Certain dilemma is inherent in relational accounts of space and time. If any objects endure through change, then temporal elements other than relations are required to describe them. If, on the other hand, no objects endure through change, no permanent reference system is available in terms of which to define the 'same place' at different times. An argument which, by exploiting this latter difficulty, attempts to show that 'objects with some endurance through time' must be accepted as fundamental is examined and found inconclusive. A sketch is then given of an alternative scheme which does allow the relevant spatial comparisons, but which does not countenance the reidentification of particulars. The discussion is intended to show that the relationist can, as indeed he must, deny the second horn of this dilemma.

M. Martin, 'The Explanatory Value of the Unconscious'

In recent years Arthur Pap and A. C. MacIntyre have argued that Freud's theory of the unconscious is not explanatory in psychoanalytic theory. But a close examination of Pap's and MacIntyre's arguments reveals that they are invalid. If one wishes to show that the theory of the unconscious is unexplanatory, different arguments will be necessary.

Philosophy of Science, 1964, 31, No. 3

W. Sellars, 'Induction as Vindication'

This paper attempts to give a rounded, if schematic, account of the concept of probability. The central concern will be to clarify the sense in which law-like statements (including 'statistical' law-like statements) are made probable by observational data which, in a sense equally demanding analysis, 'accord' with them.

R. Rosen, 'The Gibbs' Paradox and the Distinguishability of Physical Systems'

The Gibbs' Paradox is commonly explained by invoking some type of 'principle of indistinguishability', which asserts that the interchange of identical particles is not a real physical event, i.e. is operationally meaningless. However, if this principle is to provide a satisfactory resolution of the Paradox, it must be operationally possible to determine whether, in fact, two given systems are identical or not. That is, the assertion that the Gibbs's Paradox is resolvable by an indistinguishability principle actually is an assertion that we can in principle possess a complete set of effective procedures for determining the identity or non-identity of arbitrary physical systems. We show that, in rather general situations, an assertion of this type is not well founded. It is further pointed out that a failure to recognise an incomplete set of 'sameness criteria' can lead to serious blunders in physics and in biology.

M. S. Krause, 'Disconfirmative Results and Prior Commitments'

One style of conducting scientific research involves the application of theoretical propositions to an empirical domain under investigation. The investigator plans his research in terms of the theoretical propositions and states his findings in these terms as well. We wish to examine not the merits of this style of research but a difficulty inherent in it. The difficulty concerns what is to be done when disconfirmative results are produced by some discrete study or experiment in the course of such research.

R. W. Workman, 'What Makes an Explanation'

Newtonian theory has usually been accepted as a paradigm example of an explanation. There are two widely known analyses of what makes it so. According to one analysis, the deductive and predictive nature of the theory is what counts. The second analysis emphasises the ability of the theory to connect widely different events and laws. The present paper proposes a third analysis stressing three characteristics. (1) The explanation includes a description which is in part of something unobserved. (2) The description is true in the sense of corresponding to the facts. (3) Through the description, the explanation confers 'naturalness' upon the thing explained.

R. Rogers, 'Mathematical and Philosophical Analyses'

To a very significant extent mathematics is concept analysis, and although the analysis of mathematical concepts is in a number of ways different from the analysis of philosophic concepts, the similarities between these two types of concept analyses are as important and far reaching as the differences. Because mathematics and philosophy are each concerned with the analysis of concepts, they are much more like one another epistemologically than is often recognised. By virtue of the fact that mathematics, like philosophy, is concerned with the analysis of concepts, many at least of the propositions advanced within it are inherently revisable, and do not possess the kind of certainty the rationalists ascribed to them.

C. W. Berenda, 'On the Cosmological Indeterminacy Principle of McCrae'

A recent proposal by Dr W. H. McCrae, cosmologist and mathematician, to the effect that decisions between such cosmogonies as those of Hoyle and of Gamow are experimentally impossible by virtue of a general cosmological indeterminacy principle, is here examined and elaborated upon. If McCrae's principle is accepted, we will have moved a long way toward Kant's conclusion that a logically consistent understanding of an underlying 'ultimate reality' is outside the confines of science.

J. J. Fleming, 'Sub-Quantum Entities'

The possibility of a sub-quantum mechanical realm has been investigated in recent years by DeBroglie, Bohm, Vigier, and others. By restricting attention to quanta of light energy a theoretical expression is presented for a sub-quantum or micro-photon which has proven to be testable. It is possible by this development to bridge the particle-wave duality wherein these two conceptions become arithmetically commensurate with one another. The theoretical expression derived for the micro-photon also leads to a different interpretative approach to Planck's constant as well as the appearance of a new unit or element of mass.

Methodos, 1961, **13**, Nos. 49/50

F. Henze, 'The Concept of Relevance'

How does 'relevance' function in determining which facts bear on an issue, and what does it mean to say that they do? *Definitions* of 'relevance' may be shown to be circular, trivial, arbitrary, or merely persuasive, despite beliefs that relevance is sometimes a rational consideration. Five different senses of 'relevance' are described analytic, causal, psychological, normative, and performative. An analogy between admitting and judging evidence in court, and assembling facts and inferring a conclusion from them in art criticism or morals is developed.

M. Maruyama, 'Algebra of Interpersonal Interaction'

An algebraic model of sequential patterns of misunderstandings between two persons P and Q is constructed. Given this model, this paper examines the following: (1) The sequential patterns of misunderstandings, especially loops and the relationships between the loops, computed when the coding functions, decoding functions and internal logics are given; (2) The reconstruction of the internal logics when the sequential patterns of interaction are given; (3) The relationships between the loop and the algebraic ideal with respect to the internal logics.

H. Khatchadourian, 'Natural Objects and Common Names'

The paper shows that the names of 'natural objects' (unlike names of artifacts, etc.) fall into a number of different sub-classes; that some names of 'natural objects' involve one *type* of defining or criterion feature, while other names involve a different type of defining or criterion feature; and so on. The analysis, if correct, shows in effect that both the traditional 'essentialist' account and the Wittgensteinian, non-essentialist account (in *Philosophical Investigations*) are partly right and partly wrong, as far as the ordinary uses of names of natural objects are concerned.

The second part of the paper is devoted to a discussion of some of the major uses of the words 'good' and 'poor' in relation to names of natural things.

V. E. Giuliano, P. E. Jones, R. F. Meyer, G. E. Kimball and B. A. Stein, 'A Gestalt Method for Pattern Recognition'

A Gestalt method for the automatic recognition of two-dimensional patterns is described. An example of one applciation of the method is discussed, as are the results of a computer simulation of a character reader based on the application.

C. Baffico and L. Rebolia, 'Programmatore/analizzatore probabilistico automatico: risultati sperimentali'

The behaviour of the Automatic Probabilistic Programmer/Analyzer for pattern recognition has been investigated by elaborating the output punchcards for the IBM 650. Its performance has been studied by means of variations in the number of A-units, and by numerous examples given to the machine for learning. The percentage of symbols recognised, in experiments of two-class discrimination, is given.

M. Pacak, 'Syntagmatic Limits of Morphological Sets'

The paper deals with admissible sequences of words in the Czech language. The study is formal, not semantic.

J. A. Moyne, 'A Flow-charting Approach to IC Analysis'

A flow-chart is defined as a digrammatic solution for a problem to be processed by an electronic computer. Such solution is proposed for immediate constituent analysis with reference to the ultimate goal of an automatic grammar for generating sentences in a given language. A model is now proposed for establishing the nuclei of the noun and verb phrases and the endocentric structures for each. This model is applied to a flow-charting approach in which the number of the attributes of a nucleus need not be known or predetermined, because these attributes are tested and recorded by an automatic computer device known as 'address modification'. The device is generalised to recognise the subject, predicate, and object in an English sentence and to do further analysis down to the 'word' cut level by a computer.

RECENT PUBLICATIONS ON THE PHILOSOPHY OF SCIENCE

(a) BOOKS RECEIVED FOR REVIEW

Agazzi, E., *La Logica Simbolica*, La Scuola, Brascia, 1964, pp. 396, L. 1800

Albritton, Jr., C. C. (Ed.), *The Fabric of Geology*, Freeman, Cooper, California, 1964, pp. ix+374

Angell, R. B., *Reasoning and Logic*, Appleton-Century-Crofts, New York, 1964, pp. xiv+625, $7.95

Bar-Hillel, Y., *Language and Information*, Addison-Wesley Publishing Company, Inc. London and The Jerusalem Academic Press, 1964, pp. x+388

Baumrin, B. (Ed.), *Philosophy of Science*, Delaware Seminar 2, John Wiley, London, 1964, pp. xviii+551, 109s.

Benardete, J. A., *Infinity, An Essay in Metaphysics*, Oxford, 1964, pp. xi+289, 45s.

Berkner, L. V., '*The Scientific Age—The Impact of Science on Society*', Yale University Press, 1965, pp. xviii+137, 30s.

Bolacchi, G., *Teoria Delle Classi Sociali*, Edizioni Ricerche, 1963, pp. xi+256, L. 3.000

Bridgman, P. W., *The Nature of Physical Theory*, John Wiley, 1964, pp. vi+136, 12s.

Browning, D., *Act and Agent*, Atlantic Printers, 1964, pp. vii+134, $3.95

Clagett, M., *Archimedes in the Middle Ages—Vol. I, The Arabo-Latin Tradition*, Wisconsin Press, 1964, pp. xxix+720, $12.00

Crombie, A. C. and Hoskin, M. A. (Eds.), *History of Science*, W. Heffer, Cambridge, Vol. 3, 1964, pp. vi+166, 30s.

Crombie, I. M., *Plato the Midwife's Apprentice*, Routledge & Kegan Paul, London, 1964, pp. x+195, 25s.

Dawson, R. (Ed.), *The Legacy of China*, Oxford University Press, 1964, pp. xix+392, 30s.

Diderot-Goethe, *Ensayo sobre la pintura y Comentario al ensayo sobre la pintura*, 1964, Argentina, pp. 169

Dopp, J., *Notions De Logique Formelle*, Editions Nauwelaerts, 1965, p. 302, 300 FB.

Epicurus (Translated by Russel M. Geer), *Letters Principal Doctrines and Vatican Sayings*, Bobbs-Merrill, New York, 1964, pp. xli+101, $1.25

Farrington, B., *The Philosophy of Francis Bacon*, Liverpool University Press, 1964, pp. 139, 25s.

Feys, R. (Ed.), *Modal Logics*, Editions Nauwelaerts, 1965, pp. xiv+219, 350 FB.

Gregg, J. R. and Harris, F. T. C. (Ed.), *Form and Strategy in Science*, Reidel, 1964, pp. vii+476, Dfl. 38

Grünbaum, A., *Philosophical Problems of Space and Time*, Routledge & Kegan Paul, London, 1964, pp. 448, 70s.

Halloran, J. D., *The Effects of Mass Communication*, University of Leicester, 1964, pp. 83, 7s. 6d.

Handy, R., *Methodology of the Behavioural Sciences*, Charles C. Thomas, 1964, pp. xi+182, $6.75

Harré, R., *Matter and Method*, Macmillan, London, 1964, pp. x+124, 16s.

Hawkins, D., *The Language of Nature*, W. H. Freeman, London, 1964, pp. xii+372, 54s.

RECENT PUBLICATIONS

Hotopf, W. H. N., *Language, Thought and Comprehension*, Routledge & Kegan Paul, London, 1965, p. x+349, 40s.

Joos, M., *The English Verb—Form & Meanings*, University of Wisconsin Press, 1964, pp. x+251, $5.00

Källén, G., *Elementary Particle Physics*, Addison-Wesley Publishing Co., Mass., 1964, pp. xiv+546

Keller, A. G., *A Theatre of Machines*, Chapman & Hall, London, 1964, pp. viii+115, 35s.

Klibansky, R., Panofsky, E., and Saxl, F., *Saturn and Melancholy*, Nelson, London, 1964, pp. xviii+516, 6 gns.

Kyburg, H. E. and Smokler, H. E., *Studies in Subjective Probability*, Wiley, 1964, pp. vii+203, 36s.

Lamouche, A., *D'Une Morale de l'amour—à une sociologie de la raison*, Dunod, Paris, 1964, pp. 411, 44F.

Lassere, F., *The Birth of Mathematics in the Age of Plato*, Hutchinson, London, 1964, pp. 191, 30s.

Loeb, J., *The Mechanist Conception of Life*, The Belknap Press of Harvard University Press, Cambridge, Mass., 1964, pp. xlii+216, 34s.

Lonsdale, K., *I Believe* . . ., Cambridge University Press, 1964, pp. 56, 4s. 6d.

Mandelbaum, M., *Philosophy, Science and Sense Perception*, Johns Hopkins Press, Baltimore; Oxford University Press, London, 1964, pp. 262, 52s.

Marcucci, S., *L' ' Idealismo ' Scientifico Di William Whewell*, Ist. Di Filosofia, 1963, p. ix+276, L. 3000

Merleau-Ponty, M., *Phenomenology of Perception*, Routledge & Kegan Paul, London, 1962, pp. xxi+466, 56s.

> In this country French existentialism usually means Sartre, but in France Merleau-Ponty was widely regarded as the most accomplished philosopher. This book is perhaps the chief existentialist work to come from the French school.

Moore, E. C. and Robin, R. S., (Ed.), *Studies in the Philosophy of Charles Sanders Peirce*, University of Massachusetts Press, 1964, pp. xiii+225, $8.50

Morris, C., *Signification and Significance*, M.I.T. Press, 1964, pp. xii+99, 45s.

Nef, J. U., *The Conquest of the Material World*, University of Chicago Press, 1965, pp. xii+408, $8.95

Nelson, L., *System der philosophischen Rechtslehre und Politik*, Verlag Öffentliches Leben, 1964, pp. xiv+566, 38 DM

Ottaviano, C., *La Tragicita Del Reale ovvero La Maliconia Delle Cose*, Cedem, Padova, 1964, pp. xvii+775, L. 7.500

Pavese, R., *Fisicia E Metafisica*, Cedam, Padova, 1965, p. 262, L. 2500

Phenix, P. H., *Man and his becomings*, Rutgers, 1964, pp. xii+115, $3.50

Radhakrishnan, Dr. S., *Souvenir Volume—76 Articles by Eminent Writers of the World*, Darshana International, India, 1964, pp. x+600, Rs. 30.00, $8.00, £2 10s.

Ratner, S. and Altman, J., with Wheeler, J. E. (Ed.), *John Dewey & Arthur F. Bentley, A Philosophical Correspondence 1932-1951*, Rutgers University Press, New Brunswick, N.J., U.S.A., 1964, pp. xii+737, $12.50

Ritterbush, P. C., *Overtures to Biology*, Yale University Press, London, 1964, pp. xi+287, 56s.

Romer, A., *The Discovery of Radioactivity and Transmutation*, Dover Publications, New York, 1964, pp. xiii+249, 13s. 6d.

Sarton, G., *A History of Science*, Wiley, 1964, pp. xxvi+646, $2.65

Schilpp, P. A. (Ed.), *The Philosophy of Rudolf Carnap*, Open Court Publishing Co., Lasalle, Illinois, 1963, pp. xvi+1088, $15.00

Schon, D. A., *Displacement of Concepts*, Tavistock Publications, London, 1963, pp. xiii+208, 30s.

This book puts forward a theory of metaphor to explain discovery both by individuals and in the history of ideas.

Schrödinger, E., *My View of the World*, Cambridge University Press, 1964, pp. 110, 18s.

Shwayder, D. S., *The Stratification of Behaviour*, Routledge & Kegan Paul, London, 1965, pp. xvi+411, 56s.

Suppes, P. and Hill, S., *First Course in Mathematical Logic*, Blaisdell, 1964, pp. xii+274

Taton, E. (Ed.), *The Beginnings of Modern Science*, Thames & Hudson, 1964, pp. xx+665, £6 6s.

Toulmin, S. and Goodfield, J., *Discovery of Time*, Hutchinson, London, 1965, pp. 280, 35s.

Valery, P., *Aesthetics, The Collected Works of Paul Valery*—Edited by Jackson Mathews, Routledge & Kegan Paul, 1964, pp. xiv+307, 40s.

Vivian, F., *Human Freedom and Responsibility*, Chatto & Windus, London, 1964, pp. 181, 21s.

Wilkins, L. T., *Social Deviance*, Tavistock Publications, London, 1964, pp. xv+311, 45s.

Wyburn, G. M., Pickford, R. W., and Hirst, R. J., *Human Senses and Perception*, Oliver & Boyd, Edinburgh, 1964, pp. xii+340, 45s.

(*b*) ARTICLES RECEIVED

General

Apostel, L., ‘ Le problème formel des classifications empiriques ’, *Gembloux J. Duculot*, 1963, 157-230

Bose, Satyendra N., ‘ Man in Scientific Age ’, *Ann. Jap. Ass. Phil. Sc.*, 1964, **2**, 50-54

Cooper, N., ‘ ‘ The Aims of Science ’, *Phil. Quart.*, 1964, **14**, 328-333

Delogrammatik, M., ‘ La méthodologie du pluralisme et la crise de la théorie sociale bourgeoise contemporaine ’, *Filos. Nauki*, 1963, **2**, 150-156

Dretske, F. I., ‘ Observational terms ’, *Philos. Rev.*, 1964, **73**, 25-42

Fartos Martinez, M., ‘ Analitica de los principios. Leyes puras del pensar que condicionan una ciencia pura “ a priori ” ’, *G. Metafis*, 1963, **17**, 407-411

Hutten, E. H., ‘ Uncertainty and Information ’, *Scientia*, 1964, **58**, 1-8

Kyogoku, J., ‘ Science as a Culture in Contemporary Japan ’, *Ann. Jap. Ass. Phil. Sci.*, 1964, **2**, 27-30

Marciszewski, W., ‘ O zaleznosci stopnia konfirmacji od stponia eksplikacji ’, *Roczn, filoz, Pol.*, 1961, **3**, 98-107

Markiewicz, H., ‘ Scientific Laws in the Science of Literature ’, *Stud. Filo.*, 1964, **2**, 3-38

Mindel, J., ‘ Uses of Metaphor: Henry Adams on Science ’, *J. Hist. of Ideas*, 1965, **26**, 89-102

Motchalov, I. I., ' Vernadsky sur la logique et la méthodologie de la création scientifique ', *Vopr. Filos.*, 1963, **17**, 106–117

Nagel, E., ' The Nature and Aim of Science ', *Voice of America Forum Lectures, Philosophy of Science Series No.* **1**, U.S. Information Agency, Washington, 25, D.C.

Needham, J., ' Science and Society in East and West ', *Centaurus*, 1964, **10**, 174–197

Perelman, Ch., ' Reflexions philosophiques sur la classification ' *Gembloux, J. Duculot*, S.A., 1963, 231–236

Rocca, M., ' Ultime notisie dal settore scientifico ', *Sophia*, 1964, **32**, 48–53

Sadovski, V. N., ' Les problèmes de la méthodologie des sciences déductives ', *Vopr. Filos.*, 1963, **17**, 63–75

Sawada, Nobushige, ' Logic, Cybernetics and Philosophy ', *Ann. Jap. Ass. Phil. Sc.*, 1964, **2**, 55–60

Sellars, W., ' Abstract entities ', *Rev. Metaphys.*, 1963, **16**, 627–671

Shapere, D., ' The Structure of Scientific Revolutions ', *Phil. Rev.*, 1964, 73, 383–394

Shah, A. B., ' The Philosophy of Science ', *Ind. J. Phil.*, 1962, **3**, 256–268

Stegmueller, W., ' NeurePublikationen sur logisch-mathematischen Grundlagenforschung ', *Kant-Stud.*, 1963, **54**, 317–335

Stefanov, N., ' Le rôle de la mêthode dans la recherche scientifique ', *Filos. Misl., Bulg.*, 1963, **19**, 16–28

Tondl, L., ' O nekterých základnich pojmech kybernetiky ', *Otaz. marxist. Filoz.*, 1963, **18**, 201–215

Vesey, H. N. A., ' Knowledge without observation ', *Philos. Rev.*, 1963, **72**, 198–212

Yamanouchi, T., ' Analysis and Synthesis in Science ', *Ann. Jap. Ass. Phil. Sc.*, 1964, **2**, **18–20**

Logic and Methodology

Agassi, J., ' Empiricism and Inductivism ', *Philos. Stud.*, 1963, **14**, 85–86

Anton, P., ' The Status of the Empirical Principle ', *Ann. Jap. Ass. Phil. Sc.*, 1964, **2**, 21–26

Apostolova, I., ' Cause et relation causale ', *Filos, Misl.*, 1962, **18**, 16–27

Ashby, W. R., ' Induction, Prediction and Decision-making in Cybernetic Systems ', *Induction*, 1963, 55–66

Black, M., ' The Justification of Induction ', *Voice of America Forum Lectures, Philosophy of Science Series No. 15*, U.S. Information Agency, Washington, 25, D.C.

Harman, G. H., ' The Inference to the Best Explanation ', *Phil. Rev.*, 1965, **74**, 88–95

Hempel, C. G., ' Scientific Explanation ', *Voice of America Forum Lectures, Philosophy of Science Series No. 11*, U.S. Information Agency, Washington, 25, D.C.

Henkin, L., ' Completeness ', *Voice of America Forum Lectures, Philosophy of Science Series No. 5*, U.S. Information Agency, Washington, 25, D.C.

Henkin, L., ' Truth and Provability ', *Voice of America Forum Lectures, Philosophy of Science Series No. 4*, U.S. Information Agency, Washington, 25, D.C.

Hesse, M., ' A new look at scientific explanation ', *Rev. Metaphys.*, 1963, **17**, 98–108

Hesse, M., ' Induction and Theory-Structure ', *Rev. Metaphys.*, 1964, **18**, 109–122

Holton, G., ' Presupposition in the Construction of Theories ', *Science as a Cultural Force* (Ed.), Woolf, 1964, 77–108

Kleene, S. C., ' Computability ', *Voice of America Forum Lectures, Philosophy of Science Series No.* **6**, U.S. Information Agency, Washington, 25, D.C.

Kotarbińska, J., ' The Dispute Concerning the Limitations of Application of Logical Methods ', *Stud. Filo.*, 1964, **3**, 25–48

Krajewski, W., ' The Essence of Causal Relationship ', *Stud. Filo.*, 1964, **36**, 75–99

Levi, I., ' Utility and Acceptance of Hypotheses ', *Voice of America Forum Lectures, Philosophy of Science Series No. 2*, U.S. Information Agency, Washington, 25, D.C.

Madzierski, S., ' Pojecie determinizmu fizykalnego ', *Roczn. filoz.*, 1961, **9**, 43–63

Mejbaum, W., ' Falsification of Hypotheses ', *Stud. Filo.*, 1964, **9**, 49–58

Mikhova, N., ' Le probleme da la causalite ', *Filos. Misl.*, 1962, **18**, 118–128

Mortimer, H., ' Probabilistic definition as exemplified by a definition of genotype ', *Studia Logica*, 1964, **15**, 160–162

Pró, D., ' Estructura de las teorias cientificas ', *Philosophia Argent.*, 1963, **27**, 50–67

Putnam, H., ' Probability and Confirmation ', *Voice of America Forum Lectures, Philosophy of Science Series No. 10*, U.S. Information Agency, Washington, 25, D.C.

Quine, W. V., ' Necessary Truth ', *Voice of America Forum Lectures, Philosophy of Science Series No. 7*, U.S. Information Agency, Washington, 25, D.C.

Randall, Jr., J. H., ' J. S. Mill's Empiricism ', *J. Hist. Ideas*, 1965, **26**, 59–88

Rudnicki, K., ' On the Morphological Method ', *Stud. Filo.*, 1964, **36**, 101–118

Sacksteder, W., ' Theories and their context ', *Philos. Phenomenol. Res.*, 1963, **24**, 48–60

Schlesinger, G., ' It is False That Overnight Everything Has Doubled in Size ', *Phil. Stud.*, 1964, **15**, 65–71

Sellars, W., ' The paradox of analysis ', *Analysis*, 1964, **24**, 84–98

Sharpe, R. A., ' The Logical Status of Natural Laws ', *Inquiry*, 1964, **7**, 414–416

Sommers, F. T., ' Truth-functional counter factuals ', *Analysis*, 1964, **24**, 120–126

Stonert, H., ' Languages and theories adequate to the ontology of scientific language ', *Studia Logica*, 1964, **15**, 76–78

Suppes, P., ' What is a Scientific Theory ', *Voice of America Forum Lectures, Philosophy of Science Series No. 8*, U.S. Information Agency, Washington, 25, D.C.

Wald, H., ' Obsesia deductiei in explicarea inductiei ', *Crec. filoz.*, 1963, **10**, 1211–1223

Weingartner, P., Problems and Requirements in Theory Rendering ', *Phil. & Phen. Res.*, 1964, **25**, 85–95

Wright, G. H., ' Practical Inference ', *Philos. Rev.*, 1963, **72**, 159–179

Epistemology

Achinstein, P., ' On the Meaning of Scientific Terms ', *J. Phil.*, 1964, **61**, 497–508

Afanasiev, V. G., ' Les principes qui régissent la classification des systèmes généraux, *Vopr. Filos.*, 1963, **17**, 31–43

Ajdukiewicz, K., ' Empiricism and the Conception of Meaning ', *Studia Filo.*, 1964, **36**, 3–14

Davidson, D., ' On Mental Concepts and Physical Concepts ', *Ann. Jap. Ass. Phil. Sci.*, 1964, **2**, 44–49

Eskenasy, H., ' Le reflet de la logique des leins de causalité objective entre phénomènes ', *Filos. Misl., Bulg.*, 1963, **19**, 69–77

Margolis, J., ' The Privacy of Sensations ', *Ratio*, 1964, **6**, 147–153

Mixailova, L., ' Le caractere des représentations dans la science contemporaine ', *Filos. Nauki*, 1963, **2**, 48–55

Nelson, J. O., ' On the Impossibility of Sense Data ', *Ratio*, 1964, **6**, 144–146

RECENT PUBLICATIONS

Sayre, K. M., 'On Disagreement About Perception', *Inquiry*, 1964, **7**, 143-162

Russell Hanson, N., 'Observation and Interpretation', *Voice of America Forum Lectures, Philosophy of Science Series No. 9*, U.S. Information Agency, Washington, 25, D.C.

Slote, M. A., 'An Empirical Basis for Psychological Realism', *J. Phil.*, 1964, **61**, 530-537

Taylor, D. M., 'The Location of Pain', *Phil. Quart.*, 1965, **15**, 53-62

Valt, L. O., 'La corrélation entre la structure et les éléments', *Vopr. Filos.*, 1963, **17**, 44-53

Weiss, P., 'Our Knowledge of What is Real', *Rev. Metaphys.*, 1964, **18**, 1-22

Wojciechowski, J. A., 'Do scientific laws give a true image of reality?', *Proc. amer. cathol. philos assoc.*, 1963, **37**, 206-211

Mathematics

Benacerraf, P., 'What Numbers Could Not Be', *Phil. Rev.*, 1965, **74**, 47-73

Bochner, S., 'Why Mathematics Grows', *J. Hist. Ideas*, 1965, **26**, 3-24

Bradley, R., 'Geometry and necessary truth', *Philos. Rev.*, 1964, **73**, 59-75

Chihara, C. S., 'On the Possibility of Completing an Infinite Process', *Phil. Rev.*, 1965, **74**, 74-87

Goldstein, B. R., 'A Treatise on Number Theory from a Tenth Century Arabic Source', *Centaurus*, 1964, **10**, 129-160

Goodman, N., 'Science and Simplicity', *Voice of America Forum Lectures, Philosophy of Science Series No. 16*, U.S. Information Agency, Washington, 25, D.C.

Grünbaum, A., 'Carnap's Views of the Foundations of Geometry', *The Philosophy of Rudolf Carnap* ed. Paul A. Schlipp, 1963, 599-684

McShane, P., 'The foundations of mathematics', *Mod. Schoolman*, 1963, **40**, 373-387

Zabeeh, F., 'Hume on Pure and Applied Geometry', *Ratio*, 1964, **6**, 185-191

Physical Sciences

Ambartsumain, V. A., 'Les problèmes de l'astronomie contemporaine et la physique du microcosme', *Vopr. Filos.*, 1963, **17**, 45-52

Amsterdamski, S., 'Various Concepts of Determinism', *Stud. Filo.*, 1964, **2**, 85-95

Bondar, J., 'Téoria organicizmu, úrovni a finitizmu súcasnej filozofii priodných vied', *Otaz. marxist. Filoz. Ceskosl.*, 1963, **18**, 215-228

Feyerabend, P., 'Problems of Microphysics', *Voice of America Forum Lectures, Philosophy of Science Series No. 17*, U.S. Information Agency, Washington, 25, D.C.

Garkavenko, R. V., 'Diversité qualitative des particules chimiques et certaines questions méthodologiques de la chimie', *Vopr. Filos*, 1963, **17**, 89-99

Grünbaum, A., 'The Anistropy of Time', *The Monist*, 1964, **48**, 219-247

Grünbaum, A., 'Is a Universal Nocturnal Expansion Falsifiable or Physically Vacuous?', *Phil. Stud.*, 1964, **15**, 71-80

Grünbaum, A., 'Space and Time', *Voice of America Forum Lectures, Philosophy of of Science Series No. 12*, U.S. Information Agency, Washington, 25, D.C.

Hanson, N. R., 'Conservation "beyond the edge"', *Aus. J. Phil.*, 1964, **42**, 227-231

Kaminski, S., 'Problem prawdy w fizyce', *Roczn. filos.*, 1961, **9**, 85-96

Kargon, R., 'The Decline of the Caloric Theory of Heat', *Centaurus*, 1964, **10**, 35-39

Knudsen, O., 'A Note on Newton's Concept of Force', *Centaurus*, 1964, **9**, 266-271

Kouleliev, K., 'La loi de l'unité et la lutte des contraires en chimie', *Filos. Misl.*, 1963, **19**, 40–51

Polikarov, A., 'Cu privire la legea cauzalitatii in fizica', *Cerc. filoz.*, 1963, **10**, 1103–1131

Schlegel, R., 'Light velocity in the interaction interpretation of relativity theory', *Rev. Metaphys.*, 1963, **17**, 286–288

Soloview, Y. I., 'Nouveaux matériaux relatifs à la lutte des chimistes et des physiciens les plus représentatifs contre la théorie énergétique d'Ostwald', *Vopr. Filos.*, 1963, **17**, 87–96

Todeschini, M., 'Relatività di Galilei e Relatività di Einstein', *Sophia*, 1964, **32**, 44–47

Zoubkov, I. F., 'Questions théoriques controversées dans les sciences de la terre', *Vop. Filos.*, 1963, **17**, 124–130

Biological Sciences

Aron, Y. H., 'Quelques questions de la dialectique de l'universel, du spécifique et le l'unique dans la biologie et le problème de l'espèce', *Filos. Nauki.*, 1963, **3**, 83–89

Beckner, M., 'Aspects of Explanation in Biological Theory', *Voice of America Forum Lectures, Philosophy of Science Series No. 13*, U.S. Information Agency, Washington, 25, D.C.

Brien, P., 'Classification, spéciation et evolution en biologie', *Gembloux, J. Duculot*, 1963, 47–60

Kiriakoff, S. G., 'Les fondements philosophiques de la systematique biologique', *Gembloux J. Duculot*, 1963, 61–88

Hasek, M., 'K dogmatismu v biologii', *Filos. Cas.*, 1963, **11**, 562–566

Long, D. C., 'The Philosophical Concept of a Human Body', *Phil. Rev.*, 1964 **73**, 321–337

Meyer-Abich, A., 'The Historico-Philosophical Background of the Modern Evolution-Biology', *Acta Biotheoretica*, 1964, **10**, pp. x+170

Outkina, N., 'Quelques observations sur la généralisation philosophique des problèmes de la biologie contemporaine', *Filos. Nauki*, 1963, **3**, 90–95

Poliansky, U. N., 'La définition du sujet et la méthode de la biophysique contemporaine', *Vopr. Filos.*, 1963, **17**, 155–158

Polikarov, A., 'Certains problèmes méthodologiques généraux des sciences naturelles', *Izvest. Inst. Filos.*, 1963, **8**, 5–36

Praskova, N., 'La question de l'appréciation philosophique de la génétique traditionnelle', *Filos. Misl.*, 1962, **18**, 129–135

Putnam, H., 'Symposium: Minds and Machines—Robots: Machines or Artificially Created Life?', *J. Phil.*, 1964, **61**, 668–690

Whyte, L. L., 'Internal Factors in Evolution', *Acta Bio. Theoretica*, 1964, **17**, 33–48

Social Sciences

Galtung, J., 'Balance of Power and the Problem of Perception', *Inquiry*, 1964, **7**, 277–294

Indow, T., 'Man as an Object of Scientific Study', *Ann. Jap. Ass. Phil. Sc.*, 1964, **2**, 36–43

Iyanaga, S., 'Mathematical Sciences and Humanities', *Ann. Jap. Ass. Phil. Sc.*, 1964, **2**, 31–35

RECENT PUBLICATIONS

Krawczyk, Z., ' The Philosophical Foundations of Edward Abramowski's Sociology ', *Stud., Filo.*, 1964, **2**, 63-84

Leguebe, A., ' L'evolution des principes de classification en anthropologie ' *Gembloux, J. Duculot*, 1963, 117-133

Morgenbesser, S., ' Psychologism and Methodological Individualism ', *Voice of America Forum Lectures, Philosophy of Science Series No.* 3, U.S. Information Agency, Washington 25, D.C.

Pawlowski, T., ' The choice of indices in the social sciences ', *Studia Logica*, 1964, **15**, 101-1-2

Scriven, M., ' An Essential Unpredictability in Human Behaviour ', *Voice of America Forum Lectures, Philosophy of Science Series No. 14*, U.S. Information Agency, Washington, 25, D.C.

Slavkov, S., ' Des rapports entre les mathématiques et les autres sciences ', *Filos. Misl.*, 1962, **18**, 3-15

Starchenko, A., ' Le problème de la vérité dans la théorie du droit, *Filos. Nauki*, 1963, **3**, 39-49

Szacka, B., ' On the historiosophy of Stanislaw Staszic ', *Stud. Filo.*, 1964, **2**, 39-62

Thompson, J. W., ' Psychological Methodology and Biological Theories of Thinking ', *Human Relations*, 1964, **17**, 251-257

Would subscribers kindly note that the Index for Volume XV of this *Journal* is to be distributed with the August 1965 number.

THE
REVIEW OF METAPHYSICS

A PHILOSOPHICAL QUARTERLY

Edited by RICHARD J. BERNSTEIN

Volume XVIII, No. 3 *Issue No.* 71 *March* 1965

ARTICLES	
John Wild	*Being, Meaning and the World*
Wilfrid Sellars	*The Identity Approach to the Mind-Body Problem*
Stanley Rosen	*The Role of Eros in Plato's Republic*
Stephan A. Erickson	*Leibniz on Essence, Existence and Creation*
CRITICAL STUDIES	
Eric D. Hirsch, Jr.	*Truth and Method in Interpretation*
Hector Neri Castaneda	*Knowledge and Certainty*
DISCUSSIONS	
John Lachs	*Self-Identity Without a Self*
John Bacon	*Entailment and the Modal Fallacy*
EXPLORATION	
Anton-Hermann Chroust	*Aristotle and the " Philosophies of the East "*
BOOKS RECEIVED	
Walter G. Emge and Staff	*Summaries and Comments*
PHILOSOPHICAL ABSTRACTS	
PROFESSORS EMERITI OF PHILOSOPHY	
ANNOUNCEMENTS	

Welch Hall, Yale University, New Haven, Conn., U.S.A.

Readers resident in Great Britain may subscribe at a special rate through our British representative, Professor Stephen Körner, Dept. of Philosophy, The University, Bristol 8, England

---

PHILOSOPHIA NATURALIS

Archiv für Naturphilosophie und die philosophischen Grenzgebiete
der exakten Wissenschaften und Wissenschaftsgebiete

Contents of Vol. VII/2

ARTICLES:

Hubner: Was heisst und zu welchem Ende studiert man Naturphilosophie?
Heyde: Das Bedeutungsverhältnis von φιλοσοφία und " Philosophie "
Freytag-Löringhoff: Zum Problem des Zufalls
Müller: Gesetz und Denken in der exakten Naturwissenschaft
Müller-Markus: Erkenntnisprobleme der Relativitätstheorie
Schorr: Der Begründungszusammenhang der Konzeption Hugo Dinglers

Single copy DM 11 ; *Subscription DM* 36 *per volume* (4 *issues)*

VERLAG ANTON HAIN KG . MEISENHEIM AM GLAN
(Deutschland)

HEFFERS

BOOKS ON
PHILOSOPHY
PSYCHOLOGY
THE HISTORY OF
SCIENCE

BRITISH & FOREIGN

●

W. HEFFER & SONS LTD.
Petty Cury, Cambridge

Telegrams and Cables:
HEFFER, CAMBRIDGE

Telephone:
CAMBRIDGE 58351

The Foundations of Metaphysics in Science

The reflection upon the world as revealed by the sciences is at least part of the scope of metaphysics. In this book the deliverances of the main sciences are reviewed and examined to elicit the general world-picture that they suggest, and the thesis is maintained that it gives little support to the kind of philosophy which developed from the earlier work of Russell, etc.

Muirhead Library of Philosophy
63*s*

ALLEN & UNWIN

Atlas Publ. & Distr. Co. Ltd., London - Stechert-Hafner, Inc., New York - Hermann & Cie., Paris - Nicola Zanichelli, Bologna - H. Bouvier u. Co., Bonn a/Rh. - Mezhdunarodnaja Kniga, Moscou - Gerold & Co., Wien - F. Machado & Cia, Porto - Libreria Editorial Augustinus, Madrid - Librairie Payot, S. A., Lausanne - The Maruzen Company, Tokyo.

"SCIENTIA"

INTERNATIONAL REVIEW OF SCIENTIFIC SYNTHESIS

A REVIEW COVERING ALL SCIENCES
(1965 - 59th year)

Scientific Committee: Chief-Editor : P. BONETTI
G. ABETTI - G. B. BONINO - P. CALDIROLA - L. CALIFANO - R. CANESTRARI - G. COLONNETTI - B. DE FINETTI - M. GORTANI - G. LEVI DELLA VIDA - G. MONTALENTI - E. PADOA - V. SOMENZI - S. TONZIG.

"SCIENTIA" is the only review of its type which: has a world wide circulation : deals with the most recent and fundamental problems of all branches of science : can boast among its contributors of the most illustrious men of science in the whole world : publishes the articles in the native language of their Authors (English, French, Italian, German, Spanish). Each issue includes a **Supplement** *containing the complete French translation of the articles which in the text are published in language other than French.*

"SCIENTIA" has therefore a very strong appeal to the scientific-minded reader all over the world.

Full details and a free back copy will be sent by applying to
"SCIENTIA" - ASSO (Como, Italy)
sending 1s. 11d., or equivalent amount in other currency, preferably in air-mail postage stamps of your country, merely to cover packing and postage. For a number of the current year, please send 20s. 5d., or equivalent amount in other currency, which will be deducted from the subscription price.

ANNUAL SUBSCRIPTION : £6. 15s. 5d.

MAN, THE UNIVERSE AND MIND

by PHILIP MORITZ

An Empirical Analysis of Perception, World and Mind Theories, illustrating the extended use of Scientific method, beyond the more consolidated frontiers of current science and into "the hinterland" of Epi-Science.

120 pages *8s. 6d.*

MARTIGAN PUBLICATIONS
Nethercot Mill, Combe Florey, Taunton, Somerset

The British Journal for the Philosophy of Science

VOLUME XVI AUGUST, 1965 No. 62

MAN, BEAST, AND PHILOSOPHICAL PSYCHOLOGY *

JOHN KING-FARLOW AND ELTON A. HALL

SOME philosophers have thought it their occupational duty to demonstrate ever afresh that we can count on full shares of God, Freedom, and Immortality. The softer professional conscience of others has dictated that at least they succeed in establishing some remarkable and unbridgeable chasm between man and the animal kingdom. Brutes must be shown to lack the differentia of Aristotelian rationality or of a Cartesian *res cogitans*; at least they must be held to want for Lockean abstract general ideas. The Soviet philosopher-psychologist L. S. Vygotsky has allowed some animals the Marxian dignity of tool-users,[1] but he foresaw no intellectually significant role for them in the class struggle.

It is one thing for philosophers to believe in such sorts of cleavage because they infer that scientific investigation has definitively warranted their conclusions. It is quite a different matter when they leap to insist on *a priori* grounds that science must beware of whittling away at cherished barriers between man and beast. We shall divide our necessarily limited investigation of *apriorism* in recent ' philosophical psychology ' into two parts. We hope that our examination here of some analytical philosophers, men whose books we in many ways admire, may have application to other forms of anthropocentrism in other philosophical movements. Part One, ' Extreme *Apriorism*:

---

* An earlier version of this paper was read to the Seattle meetings of the American Philosophical Association in September 1964. We are indebted to the Editor and referees of this journal for several helpful suggestions.

[1] See *Thought and Language* (ed. and trans. by E. Hanfmann and G. Vakar; Cambridge, Mass., 1962, pp. 34-37

Concepts and Language', will centre around some claims made by Mr Peter Geach in his monograph *Mental Acts*, but will include cross-references to remarks with a similar ring of extreme *apriorism* in other authors. Part Two, ' Cautious *Apriorism*: Rationality and Language ', will attempt to meet more sophisticated and scientifically informed arguments in Mr Jonathan Bennett's work *Rationality*, which very recently appeared (several years after *Mental Acts*) in the same series of *Studies in Philosophical Psychology*. Unless we are badly mistaken, it would be unfair and misleading to try to answer claims like Geach's and arguments like Bennett's at the same time. For Bennett seems largely to take reasoned decisions about the way we ought to talk, if we want to talk clearly, to be prerequisite for rational inference about men and beasts; Geach seems largely to assume that fairly obvious answers are forthcoming about all such matters at once.

1 *Extreme* Apriorism: *Concepts and Language*

In order to avoid charges of extremism for our use of the epithet ' extreme ', let us begin with some of Geach's very strongest declarations of *apriorism*—declarations that urge us to turn away unhesitatingly from numerous reports by experimental psychologists.

> Many psychologists . . . would say that an animal has acquired a concept if it has learned a discriminative response to some feature of its environment. If a rat or a dog is trained to react in a certain way whenever it has a triangle shown to it . . . then they would say it has acquired a concept of *triangle*. . . . What is at issue here is not just the way the term ' concept ' is to be used, but the desirability of comparing these achievements of rats and dogs with the performances of human beings who possess a concept of *triangle*; the psychologists I am criticising want to play down the differences between human and animal performances and I want to stress them. The life of brutes lacks so much that is integral to human life that it can only be misleading to say that they have concepts like us—as misleading as it would be to say that men have tails and women lay eggs . . . or to call the noises made by brutes ' language '. Experience in training dogs to recognize triangles can be no guide in (let us say) teaching geometry.[1]

Geach goes on to accuse experimental psychologists who investigate the possibility of enhancing *concept-formation* in animals of beclouding the issue and befogging us with a 'persuasive definition'.[1] Their experiments deserve *a priori* censure from philosophers,[2] for if such psychologists do not heed philosophers' strictures 'it is only to be expected that [they] should ask Nature cross questions and she return crooked answers'.[3] It is noteworthy that *Mental Acts* belongs to a series of monographs in which philosophers have lectured psychologists and other social scientists on their grave conceptual muddles about the logic of dreaming and their myopic failure to see that genuine philosophy and genuine social studies share the same nature. Perhaps a few restricted enquiries about the wisdom of even a limited portion of Geach's work can help arouse more general qualms about a prominent brand of philosophical *hybris*.

---

[1] P. T. Geach, *Mental Acts: Their Content and Their Objects*, London, 1957, p. 17

If we speak of concepts felicitously, according to Geach, 'the central and typical applications of the term "having a concept" are those in which a man is master of a bit of linguistic usage'.[4] The paradigm case of someone's fulfilling a sufficient condition of having the concept of a *so-and-so* is 'that he should have mastered the intelligent use (including the use in made-up sentences) of a word for *so-and-so* in some language'.[5] Concepts are capacities either for using words intelligently or for performing non-verbal mental exercises which are *sufficiently like* capacities for using words.[6] Geach finds no difficulty in obtaining satisfactory criteria for saying that a chess-playing aphasiac has retained the concepts of chess, but, 'it would be hard to devise non-verbal criteria for the patient's having retained the concept of *the day after tomorrow*'.[7] Merely discriminative or recognitional capacities cannot be counted as concepts. Geach has only to relate this account of concepts to his earlier (and apparently innocuous) doctrine that *acts of judgment* presuppose concepts,[8] and he has indeed exalted man above the beasts. For if man alone has concepts, it follows, although Geach leaves us to draw the inference unaided, that man alone makes judgments. This in turn would suggest the conclusion that since no kind of non-human animals can be said to make judgments, experimental psychology cannot possibly establish what many scientists have affirmed their hope to establish: 'the essential continuity of all living forms on the mental as well as the physical side'.[9]

---

[1] p. 42

[2] p. 18

[3] p. 19

[4] p. 13

[5] p. 12

[6] p. 13

[7] p. 13

[8] p. 11

[9] C. J. Warden and L. B. Warner, 'The Sensory Capacities and Intelligence of Dogs with a Report on the Ability of the Noted Dog "Fellow" to Respond to Verbal Stimuli', *Quart. Rev. Biol.*, 1928, pp. 1-2. We shall speak of such words in future as expressing the *thesis of continuity*. Those who find such talk of the essential continuity of all living

An immediate reaction of experimental psychologists might well be that either an *apriorist* like Geach is begging factual questions and ignoring relevant data or that his victory, if it is a victory, is purely verbal. In fact it is this verbally vacuous aura which extreme *apriorists* must try to dispel. Let us raise a few questions from which we may seek to glean some points that might be urged in their defence.

I.1. *Does a linguistically oriented account of concepts so obviously fit our intuitions of what it is to have or not to have a concept that experimental psychologists must be judged to use ' concept ' in a confusingly counter-intuitive way and so to be guilty of manipulating a persuasive definition?*

Geach claims to have *history* on his side.[1] But do we have any clearcut, easily expressible intuitions of what it is and what it is not to have a concept—intuitions which tie having concepts to simple cases of using words correctly? (Compare Geach's belief that if I use ' I ' correctly, I have a concept of *self*.) ' Having a concept of *X* ' and ' not having a concept of *X* ' are such conveniently flexible grading labels that our partially competing intuitions leave us free to use them or not to use them in the following disparate ways (among others):

(*a*) We might intelligibly say that a person really hasn't a concept of *electricity* or *insanity* because although he often uses the words as people conventionally do, he is quite unable to give anything like a clear or consistent account of what the words ' electricity ' and ' insanity ' mean. Or we might say he *does* have a concept on the same grounds—depending on the grading purpose we have in mind.

(*b*) We might intelligibly deny that a person has a concept of *phlogiston, haecceitas, absolute space, existential psycho-analysis, material substance*, or *dialectical laws of history* because, although he uses the terms exactly according to the rules of a peer group fluent in such topics and although he offers fluent accounts of their meanings, we do not believe that such terms *deserve* to be counted as intelligible.

(*c*) We might intelligibly say like some psychologists that although a child uses the term ' although ' quite satisfactorily in many contexts, he has not yet formed a proper concept in connection with the term

---

forms on the mental side too ambitious or too vague may prefer to construe the ' thesis ' as follows. The study of at least some non-human animals on our own planet, together with increased knowledge of what is involved in human learning, reasoning and communication, will reveal such interesting patterns of similarity that boasts about man's intellectual eminence will seem increasingly inappropriate or even boorish. We seek here not to support the ' thesis ', but to suggest that it is not being given proper consideration.

[1] *Mental Acts*, p. 12

because he misuses it or is unable to use it when the sort of context is varied in certain tests.[1]

(*d*) We might intelligibly deny that a person has the concept of love, *gemütlichkeit*, friendship, etc., because although he uses these words according to the conventions of his linguistic community, he seems to be incapable of experiencing the appropriate emotions. (Philosophers of art and religion take note.)

(*e*) We might intelligibly deny that a person has the concept of *personal honour*, however emotionally he uses the word in grandiose attempts at self-defence, because he does not act in certain ways and ridicules others who do.

(*f*) We might intelligibly affirm that a young child has acquired a concept of *red* and *brown* and *triangle* and *circle*, etc., because without previous verbal instructions he likes to group things of particular shapes or particular colours together in his playpen. But this highly intelligible grading use of ' has a concept of ' is not repugnant to our intuitions and seems to offer comfort to both the sorts of people Geach is attacking—certain experimental psychologists and also modern philosophers like Price who emphasise abstraction and recognition in discussing concept-formation.[2]

I.2. *Could it be urged in defence of those* apriorists *who analyse concepts in terms of linguistic prowess that we need to emphasise the private mental act of judgment in order to exorcise the bogy of Behaviourism?*[3]

There is an obvious answer. If we are upset about describing humans in purely behavioural terms, why shouldn't we also be upset about doing this with certain non-human animals? Why would it be a gross overextension of analogy to attribute judgments not only to prelinguistic children but also to crafty beasts of prey like wolves or to apes in simple tool-using experiments? Notice how Geach urges that on ' psychological ' not ' logical ' grounds we should speak of judgments as presupposing concepts.[4] Yet actually it rubs even more easily with our clusters of partially competing intuitions to speak of animals as making judgments than to speak of them in certain contexts as having concepts. Geach has in effect ruled by fiat that animals lack language and cannot make judgments. Rather similarly Professor Stuart Hampshire has insisted that all animals lack language and therefore cannot properly be said to have intentions.[1] Miss G. E. M. Anscombe has gone so far as to take issue with Wittgenstein for talking of the natural *expression* of animals' intentions.[2] According to her ruling it is acceptable to speak of animals as having intentions, but not as expressing them—for they lack language and other suitably *conventional* devices. But neither of these philosophers seem to have offered reasons even as clear as Geach's for denying the propriety of saying what seems so intuitively plausible and intelligible about animals. So they can hardly be cited to *aid* Geach's case. They can be cited as symptomatic of a common malady among too many philosophers who have been discussing such issues. They appeal to our ordinary ways of talking about ' language ', ' conventions ', ' intentions ' and so on as if they could thus browbeat us into believing that we don't *really talk seriously* when we talk in ways they find philosophically distasteful—as if this were obvious!

---

[1] See Vygotsky, *Thought and Language*, chaps. v and vi

[2] See H. H. Price, *Thinking and Experience*, Cambridge, Mass., 1962, p. 35. Bennett appears to recognise the legitimacy of this use of ' concept ' at *Rationality*, pp. 20-21.

[3] Cf. P. T. Geach, *Mental Acts*, pp. 4 ff.

[4] p. 11

I.3. *Could it be convincingly argued that a very restrictive, linguistically-oriented account of concepts is desirable because it would reflect an obvious truth about the uniqueness of man as a rule-governed and rule-understanding animal?*

According to such an argument Geach would be right to emphasise the use of a word as the paradigm sufficient condition of having a concept, since the use of a word as a word involves understanding a rule and understanding rules just isn't for the beasts. Geach speaks of the ways of children who are first taught a short series of integers by rote and then, having *grasped a rule* for continuing indefinitely ' sometimes amuse themselves by counting up to large numbers '. Geach claims that nowhere in nature outside the use of human languages are there phenomena analogous to such a ' characteristic *intellectual* performance '.[3] But before assenting to any such generalisation about missing analogies one would surely want to investigate some possibly relevant cases. One might want to look at cases of what seem to be worth calling *autonomously* playful behaviour among dolphins, for example [4]; even where a dolphin has been taught a game by rather crude stimulus-response techniques, it may be glad to play the game later on according to the rules without supervision or external stimuli. How sophisticated are the games we can teach them? How able are they to construct more complicated games from simple ones? Do they behave in ways that make plausible the description, 'They deliberately improvise variations?' Such largely experimental questions rather than mere *a priori* proclamations about the meaning of 'understanding a rule' seem relevant to claims about the uniqueness of man.

---

[1] Stuart Hampshire, *Thought and Action* (London, 1959), p. 97. Cf. T. A. Long, ' Hampshire on Animals and Intentions ', *Mind*, 1963

[2] *Intention*, Oxford, 1958, p. 5

[3] *Mental Acts*, p. 31

[4] Cf. J. C. Lilly, *Man and Dolphin*, Worlds of Science Edition; New York, 1962, pp. 71-75; W. N. Kellogg, *Porpoises and Sonar*, Chicago, 1962, pp. 18-20

Mr Peter Winch in *The Idea of a Social Science* has argued rather like Geach that there must be an extremely clearcut distinction between an animal acquiring a habit and a man learning a rule. In the former case there can be no 'reflective application of a criterion': 'what counts as "always doing the same kind of thing when the word of command is uttered" is decided by a man rather than by the dog. Indeed it would be *nonsensical* to speak of the dog's doing this... a dog's acquisition of a habit does not involve *it* in any understanding of what is meant by "doing the same thing on the same kind of occasion".... The dog who balances sugar on its nose in response to its master's command has no conception of what it would be to respond differently (because it has no conception of what it is doing at all). Hence it has no alternative to what it does; it just responds to the appropriate stimulus.'[1] If a highly trained mother sheep dog were to watch one of its puppies being trained, to wag its tail when the pup performed the command correctly, to growl or grow restless when the pup went astray, would it be *nonsensical* to speak of the mother as being able to recognise whether the same sort or appropriate sort of thing was being done in response to the same command addressed to her the day before? What would Geach or Winch say of actual reports by experimental psychologists on a dog who could obey commands like 'Head on chair', 'Foot on chair', '*Other* foot', '*Other* chair', 'Do that once more'?[2] In response to the latter command 'Do that once more' the dog would invariably repeat the appropriate previous member of a set of commands numbering dozens which he was trained to perform and many of which he may have just obeyed before that previous command. 'Following a rule', 'understanding', 'identifying', 'having a conception' are of course terms for highly open-textured concepts,[3] and we may well feel it hard to say whether certain animal achievements do or do not qualify as falling under those concepts. But this hardly warrants linguistic legislation by philosophers like Winch and Geach about the 'nonsensical'

---

[1] Peter Winch, *The Idea of a Social Science and its Relation to Philosophy*, London, 1958, pp. 60–65

[2] Warden and Warner, ibid.

[3] Cf. F. Waisman, 'Verifiability', *Logic and Language*, 1st ser. Ed. A. Flew; Oxford, 1952

—legislation which would impede serious reflection on the similarities between men and animals revealed by experimental data. It certainly does not seem to warrant the claim that such legislation is simply clarification of what we meant all along.

I.4. *Is a highly restrictive analysis of concepts in terms of actual verbal behaviour required to rescue philosophical psychology from problems about counterfactuals into which it has been plunged by Ryle's dispositional flirtations with Behaviourism?*

Early in *Mental Acts* Geach makes the following comments on Ryle's frequent use of hypothetical propositions when analysing psychological statements in *The Concept of Mind* :

> We may well be reluctant, then to expound in Ryle's fashion even psychological accounts of *actual* differences in behaviour. Our reluctance should be still greater when we are invited to regard a statement that two men, whose overt behaviour was not actually different, were in different states of mind (e.g. that one checked a column of figures with his mind on the job, the other with his mind on his domestic problems) as being really a statement that the behaviour of one man *would have been* different from that of the other in hypothetical circumstances that never arose. It ought to be, but plainly is not, generally known to philosophers that the logic of counterfactual conditionals is a very ill-explored territory; no adequate formal logic for them has yet been devised, and there is an extensive literature on the thorny problems that crop up. It is really a scandal that people should count it a philosophical advance to adopt a programme of analysing ostensible categoricals into unfulfilled conditionals, like the programmes of phenomenalists with regard to 'physical-object' statements and of neo-behaviourists with regard to psychological statements.[1]

Indeed problems about the *confirmation* of statements of the forms '$Fa \supset Ga$' and '$(x)\ (Fx \supset Gx)$' have been frequently discussed by such outstanding philosophers of science as Carnap, Goodman, Hempel, and Popper. Discussions of these problems could hardly be unknown to Ryle, as Geach implies, unless Ryle did not read the contributions to *Mind*, the journal he edits. But the existence of such problems concerning confirmation has not prompted many philosophers of science simply to *dispense* with dispositional claims, universal generalisations, and other statements of the forms '$Fa \supset Ga$' and '(x) $(Fx \supset Gx)$', when 'Fa' is unfulfilled or the universal quantifier ranges over an open class.

---

[1] *Mental Acts*, pp. 6-7. Cf. G. Ryle, *The Concept of Mind*, London, 1949, pp. 86-89

Geach later claims that no similar complaints about the use of counterfactuals can be laid at his door.

> Having criticised Ryle for his reckless use of the term ' dispositions ', I may myself be criticised for introducing concepts, which are supposed to be capacities. But to say that a man has a certain concept is to say that he *can* perform, because he sometimes *does* perform, mental exercises of a specifiable sort. This way of using the modal word ' can ' is a minimal use, confined to a region where the logic of the word is as clear as possible. *Ab esse ad posse valet consequentia*—what is can be, what a man does he can do; that is clear if anything in modal logic is clear, and no more than this is involved in my talking of concepts.[1]

' What is can be.' For purposes of formal work in modal logic one may treat *possibly* as univocal and unanalysably primitive. But this will not do here: let us take some cases of what the banality ' what is can be ' might be taken to mean or endorse.

(*a*) ' If it is true that some individual g at a particular time $t^\star$ has the property P, $(Pgt^\star)$, then, whatever property P may denote, the universal proposition $(x)(t)(-Pxt)$ cannot be shown by any unpacking of P from a complex definition to be a tautology '. With such an aseptic interpretation (a variant of the Law of Identity) a logician's talk of *ab esse ad posse valet consequentia* seems safe enough whatever property P may be.

(*b*) ' If it is true that some individual g at a particular time $t^\star$ has the property P, $(Pgt^\star)$, then the universal proposition $(t)\{(t \neq t^\star) \supset -Pgt\}$ cannot be shown merely by an unpacking of P to be a tautology '. This expresses in less deceptive form than does ' What is can be ' the apparently safe intuition that whatever is the case at some time can, from a logical point of view, equally well be the case at a different time; this natural interpretation of ' What is can be ' is easily confused with (*a*), the variant of the Law of Identity, but is neither so minimally committed nor so safe. Here too, *ab esse ad posse valet consequentia*, but at least with the restriction that the predicate constant P must not denote a property with built-in-$t^\star$-uniqueness (one-time-exclusiveness). Under interpretation (*b*) we cannot move from *esse* to *posse* if P denotes a property like ' is having a bath at the exact moment $t^\star$ of Anne Boleyn's beheading for the only brief occasion of liquid immersion in one's life '. (Rather similarly, that it is noon now does not seem to allow us the inference that it could equally well be noon at any other time.) Likewise the individual constant g must not have a definition involving $t^\star$-uniqueness.

---

[1] pp. 24-25

(*c*) ' If S is P at $t^{\star}$, then it is physically possible that S be P later on '. This is another natural and intuitively-plausible-sounding interpretation of ' What is can be '. But now we have moved away from the deductive inferences (*valet consequentia*) of the modal kind Geach invoked to the realm of induction. It is, however, a very modest piece of induction under this interpretation, and inference not too likely to plunge us into consideration of counterfactuals by raising a host of ugly ' But what if . . .' sorts of questions.

(*d*) ' If Geach is now a ϕ-er, then he has the capacity to be a ϕ-er in the future.' By now the meaning is still less clear than in the case of inferring a ' physical possibility ' and the induction is positively more risky. If Geach, pursued by a bull, is at this moment a beast-soothing warbler, a militarily admirable assessor of dead ground and distance, and a silent composer of prayers of the highest literary quality, still he might be more than rash to trade in future, on the strength of what he's now doing, on his having such capacities. *Ab actu ad scientiam non semper valet argumentum.* The English *can* and the Dog Latin *posse* of Geach's tag cover an ambiguity unmasked in French: Geach needs, for the purpose of establishing concepts, to move from a certain amount of *faire* to a bit of *pouvoir* and a lot of *savoir*. But the link between *faire* and *savoir* is hardly a matter of deductive logic. If, as Geach had earlier insisted, having a concept is having a piece of ' know-how ', a disposition of a very special kind, a consideration not only of facts but of counterfactuals is surely pertinent to many cases of determining whether someone has or has not attained a concept. So Geach's account of concepts cannot be defended on the grounds that it is useful in delivering philosophy of mind from what Geach regards as the dread bogy of counterfactual conditionals in philosophical psychology. If there is such a bogy it dogs his footsteps as well as those of the Behaviourists whom he attacks.

I.5. *Are we not at least entitled to assert without further ado: ' The life of brutes lacks so much that is integral to human life . . .' ?*[1]

Many of the capacities that lead us to believe in mental similarities between man and beast are capacities acquired by long, tedious rounds of controlled stimulation from skilled humans rather than from social intercourse within their own species. But actually a philosopher like Geach obscures rather than clarifies some crucial issues with his *a priori* pontifications about animal behaviour *in general*. On the one hand we need at least roughly to distinguish animal performances *in vita*

---

[1] *Mental Acts*, pp. 12–13

naturae from performances in a controlled, ' humanised ' environment —for example, there are apes who learn to perform certain very interesting operations with binary numbers after being subjected by men to thousands of trials.[1] On the other hand we surely need to draw distinctions between the apparent capacities of different species of non-human animals, e.g. bees, dogs, dolphins, chimpanzees, etc. There is a difference between asserting that animals in their natural environments have no concepts and asserting that they *cannot* acquire concepts in any sort of environment. This ' cannot ' smacks of a strange logical necessity rather than of empirical discovery. Nor should we be blinded by lofty generalisations about ' brutes ' to the fact that non-human animals form a highly heterogeneous set warranting different sorts of expectations from those who put questions to different areas of Nature.

I.6. *Is it not a legitimate defence of an* apriorist *like Geach that his account of concepts, though strikingly narrow, compensates for this narrowness by emphasising the* digital *and grammatical character of most human language, so crucial to the development of distinctively human thought?*[2]

In languages such as English, different sorts of words have different functions, each word being understood by reference to the linguistic rules for the *range* of ways it can be employed in sentences. Geach carefully notes that there are many and varied word-skills involved in expressing judgments grammatically.[3] And Geach, as we saw, is contemptuous of those who call *any* animal sounds ' language '. But before we can generalise about animal sounds, we need far more experimental data concerning various species whom we casually speak of as communicating by sound in their natural environment. Certainly bee communication about objects at distances can be described in terms of rules and elements, though we might not wish to say that bees are *aware* of rules.[4] Maybe further research into 'natural' dolphin communication will afford good reason for calling them self-consciously grammatical animals. As yet we have insufficient information. We should also await further experimental data about animals' ' improved ' performances with and in response to voice, gestures, or other stimuli after men devise increasingly refined techniques of training. In evaluating such future data, we should be mindful of different sets of criteria for *grasp of language*. One might cite, *inter alia*, grasp of language as instanced by animals' initiation of communication vocally and by gesture; grasp of language as instanced by awaiting and obeying commands, etc.; grasp of language as instanced by behavioural expression of pleasure or displeasure at the responses of other animals to human communication.

---

[1] See C. R. Ferster ' Arithmetic Behaviour in Chimpanzees ', *Scientific American*, 210, no. 5, 1964, pp. 98-106

[2] Cf. Noam Chomsky, ' Review of Skinner's *Verbal Behaviour* ' *Language*, 1959, 35, especially pp. 54-58

[3] *Mental Acts*, pp. 12-13

[4] See Bennett's *Rationality*, London, 1964, pp. 8-21

Consider a possible experiment with a simple imperative language having a clearcut grammar or set of rules for well-formed formulae. A well-formed formula consists of three elements in the following order: (*a*) a ' verb ' of command, such as ' fetch ', ' drop ' or ' bite '; (*b*) an ' adjective ' suitable to the particular species' tested powers of sensory discrimination, e.g. ' bright ', ' dull ', ' round ', ' triangular '; (*c*) a ' noun ' like ' basket ', ' bone ' or ' stick '. The ' words ' may or may not be transmitted orally. Now is it possible for a particular species of animal to master the grammar of such a simple command-language so that the animal will obey all and *only* well-formed formulae, refusing to respond to such misformulations as ' Basket ', ' Bring basket brown ', ' Basket bring brown ', and others that violate the grammatical rules? Further, can the animal be trained to grasp by abstraction that a new sound (or sign) is a ' noun '-sign, ' adjective '-sign or ' verb '-sign? Can the animal express (by *any* means) recognition of other animals' violation of or conformity to the rules of the command-language?

Striking success in getting the desired results from a species would still raise conceptual questions because of the already emphasised open-texture of certain concepts like ' recognising a rule '. Yet the concepts are neither so hopelessly open-textured nor so completely defined (as Geach and Winch assume) as to bar the quest for experimental data in order to answer such questions more rationally. Although the answers might involve value judgments concerning the *desirability* of drawing comparisons and altering the texture of our concepts, so long as the answers result from open interest in the discovery of possible similarities, the answers stand a chance of being rational. *A priori* answers like those of Geach and Winch that would limit our interest in experimentation stand no such chance.

2 *Modest* Apriorism: *Rationality and Language*

A more modest form of *apriorism* concerning the thesis of continuity between men and animals is evinced by Mr Jonathan Bennett, in his very recent work *Rationality*. Certainly *apriorism* is pronounced in the following statements of his aims, especially pronounced in the way a *definition* is used to prejudge the result of ' enquiry ' in flagrant violation of Lesniewski's very desirable criterion of *non-creativity* for acceptable definitions.[1]

> Now, I use ' rationality ' to mean ' whatever it is that humans possess which marks them off, in respect of intellectual capacity, sharply and importantly from all other known species '. My analysis of rationality, then, will explore the content of the true belief that human beings are on a certain intellectual eminence compared with other terrestrial creatures: it will thus include an enquiry into the truth of the belief that this eminence is surrounded by a logical precipice and not just by a gentle slope which happens to be unoccupied, and into the truth of the belief that the eminence is essentially connected with human linguistic ability.[2]

On the other hand, Bennett's *apriorism* is modest indeed compared to that of authors earlier discussed: he refers frequently to experimental data concerning bees and apes; he attempts at length to argue about what is, strictly speaking, proper usage of expressions like ' intelligent ', ' rational ', ' symbolic ', ' rule-guided ', ' reasons ', and ' language '; he refuses ' to swallow in such a greedy and undigested lump the Kant-Wittgenstein stress upon language that one sees it as just obvious that there cannot be judgments where there is no language '[3]; for the most part he uses the term ' concept ' and its associates ' notion ' and ' idea ' with the pragmatic flexibility of ordinary language and without begging serious questions about their proper application as Geach does. Bennett's major strategy in his book may well interest many philosophers, biologists, linguists and psychologists: he describes the behaviour of honeybees as reported by K. von Frisch,[4] creatures that one may speak of loosely ' in an informal, experimental report ' as having a *language*; he then tries to isolate some crucial features of human *linguistic* behaviour that are lacking in apian dances; he seeks to bring out what the really essential differences are by considering what additional forms of behaviour honeybees would have to exhibit before they could be strictly said to be rational animals with a language. Bennett ingeniously enriches the hypothetical honeybees' observable communicative behaviour in gradual stages whereby they advance to the level of

---

[1] See P. Suppes, *Introduction to Logic*, Princeton, 1955, pp. 152-160

[2] *Rationality*, p. 5 [3] p. 2

[4] Certain claims of von Frisch and Bennett about actual bee communication—notably with regard to the visibility of dances in the hive—may need to be revised in view of the researches of our colleague, Dr A. M. Wenner. See his ' Sound Communication in Honeybees ', *Scientific American*, 210, no. 4, 1964, pp. 116-124.

'intelligence' and then 'rationality'. He offers complex behavioural criteria which would make it plausible to attribute to their 'language' first both assertions *and denials*, and then '(1) statements with two distinct tenses, (2) two distinct subjects of discourse, and (3) means for stating their reasons for any particular affirmation or denial,'[1] and finally a device whereby they can deny that a reason-statement warrants particular affirmation, even when both reason-statement and affirmation are true. According to Bennett such enrichment is needed before we can properly say that the honeybees' communicative behaviour is *rule-governed* rather than merely *regular*, or that their behaviour is at least partially explainable in terms of *reasons* rather than merely in terms of a *simple stimulus-response pattern of causes.* Without quite so much enrichment it might be proper for us to speak of them as having a *language*, but not to say that they have the linguistic capacities required for being *rational.* Bennett hopes to have thus disclosed from what is all too familiar (rather than from esoteric data) a 'logical precipice' or 'great gulf' between characteristically human behaviour and that of all other known animals.

> With the introduction of universal and dated statements, there are canons of appropriateness whereby we can ask whether an apian act is appropriate not to that which is particular and present to the bee but rather to that which is particular and past or to that which is not particular at all but universal. This is what generalising and talking about the past have in common: they are both departures from that which is present and particular. This common feature is what links them with rationality. The idea of rationality is that of the ability, given certain present and particular data, to unite or relate them with other data in certain appropriate ways . . . only in a language is it possible to register theoretical knowledge, that is, knowledge-that-*p* which is not cast in the form of knowledge-how-to-get *x* . . . there is a great gulf between the sorts of competence which are expressed in dated and universal statements and the sorts which are not.[2]

Bennett's chain of argument is worth meeting at different levels; we shall be content with examining two. First, at a more 'factual' level, it is worth asking: even if we wished to accept Bennett's seemingly very stringent requirements for having reasons, having language, following rules and being rational—wished to accept them as *the proper* requirements—would we be entitled to conclude that man is unique in his rationality? It strikes us as little less than amazing that Bennett should have so much to say about bees, dogs, and apes, but nothing to say about more relevant current research concerning what are apparently highly intelligent forms of marine life, dolphins and whales. For example, scientists have expressed a keen conceptual interest in the report that a previously unsuspecting group of killer-whales was able to learn quickly after a 'communication' from one or a few of its members that all and only those fishing boats differentiated by a small distinctive feature offered danger to their kind. Once so 'warned' the killer-whales carefully avoided the armed craft but operated without hesitation among the others.[1] It is by no means certain that the behaviour of killer-whales or of their vastly more docile, observable and humanly educable relatives, the dolphins, may not make them plausible candidates for rationality even by Bennett's criteria. *A posteriori* research by marine biologists and others may help us to settle this question within a decade or so.

---

[1] *Rationality*, p. 51

[2] pp. 85, 88, and 97–98

Bennett's chain of argument is also worth challenging at another, more 'philosophical' level. He brilliantly offers a set of behavioural criteria which would make it extremely plausible to attribute tensed judgments, universal judgments, and '*p* because *q*' judgments to his hypothetical honeybees. One is likely to be so dazzled by this *tour de force* that one forgets the possibility that scientists will discover shades of increasingly complex behaviour, for example among the more promising forms of marine life, that will make it seem progressively more, but perhaps not quite completely plausible to attribute analogies of such judgment forms to a spectrum of animal species.[2] Thus, even if one postulated 'having property ϕ' and 'not having property ϕ' as the all-determining criteria for being rational or non-rational, one would not lose the option of saying (in the face of experimental data) that there appears to be a curious sort of continuum between species that behave clearly unlike men (who have property ϕ paradigmatically) and species that behave as if it very well might be reasonable to ascribe ϕ or something extremely ϕ-like to them. Is there any clear line of demarcation between animals whose behaviour is describable in terms of causal regularities alone and animals whose behaviour is describable in terms of following rules and having reasons? If it is argued in Bennett's favour that having a certain sort of language is one of the crucial notions for establishing this line, we reply that this criterion will not do for Bennett, since *following rules* and *having reasons* seem to be among the crucial notions for establishing that a species has a full-blown, sufficiently sophisticated language. What is to count as a fatally *simple* stimulus-response pattern of causal explanation? What is the allegedly clear distinction of a *sufficiently complex* pattern? Is it to be the having of a logically sophisticated language? But here again we seem to be moving in a vicious circle. We seem to have too many primitive or pseudo-primitive terms, each needing the others for explication and each too ' open-textured ' to supply the decision procedure required for a ' logical precipice ' or a ' great gulf '. We shall try to reinforce these philosophical criticisms of Bennett by looking more closely at his use of the terms ' fitting a fact ' and ' analysis '.

---

[1] See Lilly, *Man and Dolphin*, pp. 61-62. Cf. R. B. Robinson, *Of Whales and Men*, London, 1954

[2] W. V. Quine in *Word and Object*, Boston, 1960, chap. ii shows that there are severe problems about verifying the correctness of 'radical translations' from even a *human* language that is exceedingly unlike our own. It might thus be extremely hard to decide whether a crucial connective in such a language should be rendered as 'because' rather than 'and'; there might indeed appear to be no correct answer. Bennett's optimism sometimes suggests a commitment to what Quine attacks as the myth of 'language-transcendent meanings'; thus he expects animal behaviour always to supply a sufficiently clearcut decision procedure for concluding that the communicators do or do not make judgments of a certain form that we make. Lilly's first chapter, 'The Possibility of Interspecies Communication', indicates that there may also be serious biological reasons for pessimism concerning the interpretation of many animal sounds.

Fitting facts. Bennett is at his most implausible when he speaks as if we had a very clear notion of what it is for behaviour to *fit a fact.* Nor is this matter peripheral since a crucial stage in his argument that rationality presupposes use of a fairly complex language involves the claim that only linguistic behaviour can fit facts of certain kinds. Consider the following remarks:

> *Only linguistic behaviour can be appropriate or inappropriate to that which is not both particular and present* . . . a sheep dog's manoeuvre is apt or appropriate to the fact that the sheep are approaching the cliff edge. . . . But what can the poor dog do now which either fits or fails to fit the fact that it had a swim last Wednesday? . . . The only way in which behaviour can either fit or fail to fit a fact is for it to *say* something which means that some general state of affairs obtains; and the only way in which behaviour can fit or fail to fit a fact about the past is for it to *say* something which means that something was present in the past. . . . But there is no non-linguistic way of manifesting knowledge of just some fact about one's past: the dog's retrieval of a bone manifests, apart from its desire for a bone, two things at once: its belief that it buried the bone in that place, and its belief that in general bones stay put until retrieved by the burier. Only in language is it possible to manifest a belief about

the past without at the same time manifesting beliefs of a general kind.[1]

In so far as we have any clear intuitions of what it is for behaviour to 'fit a fact' we find Bennett's line of argument rather counter-intuitive. For example:

(1) 'What can a poor dog do now which either fits or fails to fit the fact that it had a swim last Wednesday?' It is far from fantastic that some dogs or dogs could be taught by present or more sophisticated methods to behave in the following non-linguistic ways that fit such a fact.

(*a*) His trainer teaches him to go and swim once and only once every alternate day after an initial day on which his trainer immerses him in water. If last Wednesday was a day when he went and swam, or a day on which he was immersed, and if today is Friday, then the successful training results in his determinedly striving to reach a stream or river. His striving to get to water would seem to count as fitting the fact that he had a swim last Wednesday.

(*b*) We find with this dog that if a strange handler leads him to water on the Friday after a Wednesday bathe, he co-operates. But if a strange handler leads him to water on a Thursday after a Wednesday bathe, he resists immersion violently.

(*c*) We find that the dog, who has observed a duller kennel mate being similarly trained, shows agitation when he and his fellow are first immersed on Wednesday and then, after completing his Friday swim, he finds that his fellow has failed to move from the bank. He barks at the dullard and tries to nudge him off the edge. If the dullard makes for the water on the following day (Saturday), he tries to head him off.

(*d*) The duller dog begins to enter the water on a Thursday. The brighter dog howls and the dullard, after pausing and cocking his head, stops going in.

(2) 'The only way in which behaviour can fit or fail to fit a fact about the past is for it to *say* something which means that something was present in the past. . . . But there is no non-linguistic way of manifesting knowledge just of some fact about one's past.' Here again it seems easy to hypothesise a counter-example. Suppose an originally sad stray dog of extremely affectionate disposition finds a kindly master at place *p* a mile from the master's house. Whenever he passes *p* in the following months, he wags his tail and leaps furiously about; if he is

---

[1] *Rationality*, pp. 87-89

with his master, he rubs against him. But a year later his master has become alcoholic, withholds affection from the dog and frequently mistreats him. The dog no longer cavorts when passing *p* but actually lets his tail droop noticeably lower. If the dog exhibits a remarkable memory in many other respects, if he seems able to distinguish *p* from numerous very similar spots in the neighbourhood, surely it would not be infelicitous to say that his behaviour, first unusually joyous, then pronouncedly droopy, manifests knowledge of the particular fact about his past that he met his master at *p*. (Would Bennett wish to retort that such behaviour cannot but also manifest a general belief that for all *t*, *p* is identical with *p*?)

Analysis. It is unfortunate that Bennett likes to talk in a Moorean way about giving an *analysis* of difficult terms, seeming to suggest that philosophical work on X's must aim to uncover what we do really, essentially and clearly mean by ' X '. One notes the subtitle of *Rationality*, ' An Essay towards an Analysis '; he remarks on page 1: ' Kant sees himself as engaged not in conceptual analysis but in the establishment of synthetic truths of a certain sort. This is a mistake of his which I cannot go into now: I merely turn my back on it, and set in place of his synthetic conclusions the search for an adequate analysis [of " rationality "].' Such talk of ' analysis ' smacks of a strain of *apriorism* of the worst kind in contemporary linguistic philosophy—of the belief that the way we do talk of X's (the way we *really* talk, what we *basically* or *essentially* mean in talking) is *a priori* the way we *ought* to talk.[1] Given the natural interpretation of ' analysis ' in this philosophical period, it would surely have been preferable for Bennett to have selfconsciously and explicitly attempted to *prescribe* a *radical reconstruction* (for definite, specified purposes) of the notions he considers. For our ordinary usage of such terms as ' rule ', ' reason ', ' intelligence ', ' language ', ' rationality ' and ' fitting a fact ' is surely likely to be subject, as we saw with ' concept ', to a host of competing intuitions. Had Bennett explicitly attempted a radical reconstruction of such notions for specific purposes, he might have saved himself both from going too far and from not going far enough: he might have saved himself from not going far enough because he would have seen the need to explicate his terms with primitive ones that would not be so open-textured and would thus be more helpful for purposes of observation and verification; he might have saved himself from going too far because the radically reconstructed notions would clearly correspond to some of our intuitions and clearly *not* correspond to others. This in turn might lead to the realisation that there could be several useful reconstructions; such a realisation would discourage dramatic dogmatism about the discontinuity between men and various species; it would encourage us to examine more carefully the many different ways in which different species are like us and unlike us.

---

[1] His general approach seems to be strongly at variance with his mild disclaimer at p. 97 that ' the word " rationality ", however, is not my central concern.' Cf. E. Gellner, *Words and Things*, chap. ii, section 5, ' From Fact to Norm '.

Conclusion

We have tried to suggest in very limited and piecemeal ways that philosophy is pernicious when its rigidifying accounts of present notions about man become a fixed barrier to any shift in those notions or to a shift in our world-view in the face of actual and possible experimental data. The investigation of actual and potential behaviour among so-called lower animals is not at present very advanced and certainly not complete. If the philosopher is genuinely open to the discovery of new patterns of similarity and dissimilarity between man and beast, he is a better philosopher for that. Why a philosopher like Geach should feel qualified to assert that what animals can do and can be trained to do must have no bearing on man's ability to learn a discipline like arithmetic or geometry is a mystery which eludes us. Robert Thomson, a philosophically knowledgeable psychologist, has recently and wisely written:

> Human thought may have its origins in similar *kinds* of operations to those practised in a simpler way by animals. And some basic types of human skill may have their parallel lower down the scale of life. The study of animals, like the study of children or primitive peoples, may give us clues regarding what is crucial and important in the make-up of the civilised adult.[1]

Postscript: A Genetic Speculation

Having discussed these examples of *apriorism* among some current analytical writings on philosophical psychology, we set before the reader a curious historical and, indeed, *genetic historical* question which the Editor of this journal has mentioned as worthy of reflection. *Why is it, now most of us no longer believe the differences between men and animals*

---

[1] R. Thompson, *The Psychology of Thinking*, Penguin, 1959

to turn on the possession or lack of an eternal soul, that so many philosophers of language have become preoccupied with questions about other grounds for drawing an essential division? At least one of the *apriorists* discussed seems to make little secret of his belief that men have souls and brutes do not [1]; this belief may even inspire his attempts to argue for an essential division on other grounds. Concerning those *apriorists* who are not influenced by any such religious scruples—scruples a Hindu or an Orphist might consider spiritually subversive—it may be worth risking some cautious historical comparisons and generalisations.

It seems unlikely that every living spokesman for any generation which has begun to set aside its forbears' preoccupation with an after-life will wholeheartedly welcome any sober investigation of the thesis of continuity. R. M. Blake's illuminating essay, 'Natural Science in the Renaissance',[2] shows that the Renaissance Humanists, in sharp contrast with a number of Augustinians and Occamists, were among those most hostile to the development of interest in the sciences. Literary and linguistic studies were jealously upheld at science's expense. Petrarch, a notable spokesman for the Renaissance's fascination with the things of this world, is cited by Blake as especially disdainful of biologists who studied animals, since such inquiries seemed quite irrelevant to the nature of man. While Renaissance Humanists were more often opposed to painstaking theology than to all religion, it is notable that in an earlier age of far deeper religious scepticism, various Greek Sophists urged philosopher-scientists away from reflection on physics and biology towards the more timely study of man, his language and conventions.

Perhaps even more striking is the haste and passion with which our stridently atheist contemporary M. Sartre (in *L'Être et le Néant*) moves from the denial that the human mind is a Cartesian substance or spirit to the claim that therefore, since mind is not a *thing*, it must *transcend* all things, and its thought cannot be explained by anything like the natural laws we apply to physical objects.

In so far as genetic explanations are of interest to the philosopher, it seems not unfair to suggest that there are some among British and American analytical philosophers, as well as among European existentialists and Marxists, who continue to exemplify a *historical tendency* to fill the aching void created by rejection of man's supernatural importance with a militant cult of man's *natural* uniqueness. Man's uniqueness is of course to be celebrated in different tones of voice by different sorts of philosophers. Gilbert Ryle tells us in *The Concept of Mind* that man is man and not a machine. One may admire Ryle's valuable indications of dissimilarities between human behaviour and that of relatively ' primitive ' machines, yet regret there is no comparable discussion of possible similarities in the case of relatively ' sophisticated ' machines. The anthropocentric bias that Ryle's reticence seems to suggest Martin Heidegger hastens to proclaim: man must ex-sist from everything else, his consciousness should set him apart as the *shepherd of Being*. Several Iron Curtain delegates to the XIIIth International Congress of Philosophy in Mexico City (1963) sought to dispel any confusions between Dialectical Materialism and ' vulgar, metaphysical materialism.' Man's singular stature as a being with higher consciousness and sentential freedom, one heard, was at least as clearly emphasised by Marx as by any other modern philosopher.[1]

---

[1] See *Mental Acts*, pp. 17 and 121

[2] In *Theories of Scientific Method.* Ed. E. H. Madden; Seattle, 1960

Thus we sympathise with Bennett when he complains of those who embrace ' the timid and reactionary anthropocentrism which can be extracted from the worst parts of the later Wittgenstein—a chest-pounding insistence on the uniqueness of the human species '. Yet the tendency seems far too widespread among different schools of philosophy and far too deep-rooted in history for Wittgenstein to be made the major scapegoat. And the question remains whether Bennett himself has not been influenced by the tendency, or unwittingly encouraged it, or both.

University of California,
Santa Barbara

---

[1] The Russian contributions appear in English translation in *Philosophy, Science, and Man*, Moscow, 1963. See especially the contribution by Academician M. B. Mitin.

THEORETICAL MODELS*

PETER ACHINSTEIN

THE term *model* enjoys a broad range of uses in the sciences. It may refer to anything from a physical construction in a display case to an abstract set of ideas. Here I shall examine one important use of the term in physics, illustrated by examples such as the Bohr model of the atom, the billiard ball model of gases, the corpuscular model of light, the shell model of the atomic nucleus, and the free-electron model of metals. These I shall refer to as *theoretical models*,[1] or *models*, for short. It is not my claim that everything called a model in physics can be shown to be a model in the present sense. Indeed, I shall argue, these models are quite distinct from other conceptions sometimes called models. It is my aim to show that the examples I have cited, as well as many others that might have been cited, have certain common characteristics in virtue of which they are referred to as models.

Philosophers of science who have emphasised the importance of models, and have claimed that a consideration of them will illuminate the structure, interpretation, and development of scientific thinking, include a number of the above on their lists. Yet these philosophers have proceeded to make claims which, I shall maintain, tend to be misleading and do not reveal what is especially characteristic of models of the sort I shall be considering. While my concern will be only with theoretical models in physics, I believe that much of what I have to say is applicable to many, but certainly not all, conceptions also called models in biology (for example, the Crick-Watson model of the DNA molecule), psychology (models of learning behaviour), and economics (the multiplier-accelerator model of economic growth), though it is not my intention to argue for this here.

I

There are four characteristics of theoretical models I propose to emphasise.

* Received 22.i.65

[1] This term is also used by Max Black in *Models and Metaphors*, Ithaca, 1962, p. 226, though he presents a somewhat different account from the one to be offered here.

1. A theoretical model consists of a set of *assumptions* about some object or system. For example, the billiard ball model of gases is that set of assumptions according to which the molecules comprising a gas exert no forces on each other except at impact, travel in straight lines except at instant of collision, are small in size compared to the average intermolecular distance, etc. The Bohr model is that set of assumptions according to which the electron revolves about the nucleus of an atom in such a way that its orbital angular momentum is quantised as is the energy radiated or absorbed by the atom. Two points deserve notice. First, theoretical models must be distinguished from diagrams, pictures, or physical constructions which, though sometimes useful in presenting the model, are not to be identified as the model itself. Second, it is sometimes though not always true that what is called a model is also called a theory, as in the case of the Bohr model of the atom and the free-electron model of metals. This interchangeability of labels is possible since, in such cases, the terms ' model ' and ' theory ' refer to the same set of assumptions, though to call this set a model is not to suggest the same things about it as when one calls it a theory. Some of the differences, as well as the reasons why not all models are called theories and vice versa, will emerge in what follows.

2. A theoretical model describes a type of object or system by attributing to it what might be called an inner structure, composition, or mechanism, reference to which will explain various properties exhibited by that object or system. The billiard ball model ascribes a molecular structure to gases in such a way as to enable the derivation of principles relating pressure, volume, temperature, entropy, etc., of gases. The corpuscular model of light ascribes a particle composition to light that purports to explain properties of light such as reflection and refraction. The Bohr model describes the underlying mechanism of the hydrogen atom in such a way as to explain radiation of discrete wavelengths observed when hydrogen is excited. Explanation by appeal to such notions covers a number of similar (though by no means identical) kinds of analyses. The following points should help to clarify some of the ideas involved.

The distinction between a set of exhibited properties of an object and its inner structure, etc., reference to which can explain these properties, is not to be confused with the distinction between ' macro-phenomena ' and ' micro-phenomena '. Many of the properties explained by a theoretical model may themselves be on the ' micro-level ' (as in the case of atomic and nuclear models). On the other hand, the properties explained as well as the structural components of the model can be on the 'macro-level' (as in the case of 'world-models' in cosmology). Moreover, the structure depicted must always be considered with reference to those properties it serves to explain and not as something ultimate or incapable of further explanation. (Thus, the molecular structure invoked in the billiard ball model is itself explained by atomic models.) Finally, the distinction in question must not be understood simply as the distinction between properties and their explanation. For not every set of principles capable of explaining properties exhibited by an object or system need refer to its inner structure. Some explanations proceed by describing just what relationships obtain among the exhibited properties themselves. In classical thermodynamics, for example, changes in certain thermodynamic properties of matter such as pressure, volume, temperature, and energy, can be explained by changes in others, although no inner structure (such as that involved in a molecular hypothesis) is ascribed to matter. Other explanations relate properties of the object to properties of quite different objects (as the tides on the earth are explained by the gravitational attraction of the sun and moon). A theoretical model, on the other hand, will proceed by analysing the object that exhibits certain known regularities into more basic components, not just by expressing these regularities in quantitative terms or by relating the known properties to those of distinct objects or systems. Accordingly, the use of the term 'theory' is in this respect broader than that of 'model' since not all theories are designed to provide structural analyses typical of models.

3. A theoretical model is treated as an approximation useful for certain purposes. One who formulates such a model will reason as follows: 'It is useful to represent *x*'s as having such and such a structure, for then various known principles can be derived; moreover, the actual structure that *x*'s have is something like this, though quite possibly more complex (or in many cases actually known to be more complex).' For example: 'It is useful to represent gases as composed of tiny elastic spheres obeying Newtonian laws, for then we can derive the ideal gas law, diffusion and transport equations, etc.; moreover, the actual structure of gases is something like this, though more complex because attractive and repulsive intermolecular forces are not considered in this representation'. The *value* of a given model, therefore, can be judged from two different though related viewpoints: how well it serves the purposes for which it is employed, and the completeness and accuracy of the representation it proposes. Is the billiard ball model a

good one? Yes and No. It does enable a number of important relationships to be derived, though by ignoring certain intermolecular forces it fails to provide a complete or accurate representation.

The fact that a theoretical model is proposed as a way of representing the structure of an object or system for certain purposes explains why there are often alternative models in use: different representations may be employed for different purposes, for example, the billiard ball model for deriving the perfect gas law, the weakly-attracting rigid-sphere model for deriving the Van der Waal's equation, the model representing molecules as point centres of inverse power repulsion for facilitating more realistic transport calculations.[1] Here then is another point of difference between the use of the terms ' model ' and ' theory '. To propose something as a *model* of (an) *x* is to suggest it as a way of representing *x* which provides at least some approximation of the actual situation; moreover, it is to admit the possibility of alternative representations useful for different purposes. To propose something as a *theory* of (an) *x*, on the other hand, is to suggest that *x*'s *are* governed by such and such principles, not just that it is useful for certain purposes to represent *x*'s as governed by these principles or that such principles approximate those which actually obtain. Accordingly, the scientist who proposes something as a theory of (an) *x* must hold that alternative theories are to be rejected, or modified, or understood as holding only for special cases, or something of the like.

This helps to explain why certain sets of principles first called theories may later be called models. Originally proposed as reflecting the actual structure of the items in question (e.g. atoms, metals), they are now recognised as approximations, still useful for certain purposes. Such conceptions may of course continue to be referred to as theories, thereby reflecting the historical fact that they once were proposed as theories and not simply as models.

4. A theoretical model is often formulated, developed, and even named on the basis of an *analogy* between the object or system described in the model and some different object or system. This involves a comparison in which properties and principles similar in certain respects are noted. The Bohr model postulates orbiting electrons on the analogy with a planetary system. The liquid-drop model of the atomic nucleus explains nuclear fission on the analogy with the division of a liquid drop into two smaller drops. The nuclear shell model was developed by invoking an analogy between the atomic nucleus and

---

[1] See R. D. Present, *Kinetic Theory of Gases*, New York, 1958, pp. 108 ff.

extra-nuclear electron shells. This conforms with the previous point that theoretical models are intended to provide a useful representation of a system: to provide such a representation it is frequently helpful to draw an analogy between the system in question and some known system governed by laws which are understood, and to suppose that some of these laws, or similar ones, also govern the system to be described in the model. Reasoning of this sort, based as it is upon an argument from analogy, is never considered sufficient to establish the principles in question but only to suggest that they may be considered reasonable first approximations, subject to test and subsequent modification. In each case, however, the model itself can be distinguished from any analogy upon which it might have been developed. The billiard ball *model* consists of those propositions asserting that gases are composed of tiny elastic spheres, etc. The *analogy* is drawn between gases so described and perfectly elastic billiard balls.

These represent what I take to be some of the more important characteristics of theoretical models. Something must now be said concerning their *use*. Theoretical models can serve all the same functions as theories: they can be used for purposes of explanation, prediction, calculation, systematisation, derivation of laws and so forth. The difference between the use of a model and the use of a theory lies not in the type of function it can be employed to serve but in how it serves that function. Theoretical models provide explanations, etc., but ones based on possibly simplified assumptions. And this qualification must be borne in mind when comparison is made to theories. Explanation and systematisation via a theory is often said to be ' deeper ' and to provide more insight, thus reflecting the belief that the principles constituting the theory are more accurate than those of the model and take into account more of the known quantities. Why then not always employ the more accurate and complete theory? First, none may exist, as is presently the case, for example, with models of the atomic nucleus. Second, even if one does exist a theoretical model may be of considerable importance. It may afford calculations impossible or exceedingly difficult with the more complex theory. In kinetic theory calculations involving transport phenomena such as diffusion, heat conduction, and viscosity become quite difficult unless simplifying assumptions embodied in several models of intermolecular force fields are adopted. Again, by studying the consequences of deliberately simplified assumptions constituting a model the scientist may gain further insight into the more complex theory and may even be led to extend or modify it. Finally, theoretical models serve an important didactic role. It is frequently a simpler procedure to study the Bohr model before turning to quantum theory; to treat transport phenomena by assuming a billiard ball model before learning the more advanced Maxwell-Chapman transport theory; and so forth.

2

I propose now to examine a number of assumptions that are frequently made about theoretical models. One is that a model is designed to provide an interpretation for an otherwise uninterpreted formalism or calculus. Nagel, for example, speaks of a model as 'an interpretation . . . for the abstract calculus which supplies some flesh for the skeletal structure in terms of more or less familiar conceptual or visualisable materials'.[1] Hutten asserts that 'the model gives a possible interpretation to the symbols (in an equation or formula) which thereby acquire a meaning, and we can apply the equation, or formula, and test it'.[2] Braithwaite defends a similar position.[3] A second assumption is that a model is always proposed with reference to some theory, where the model has the same formal structure as this theory. Braithwaite speaks of a model as 'another interpretation of the theory's calculus',[4] and holds that 'similarity in formal structure . . . is all that is required of the relationship of model to theory'.[5] Nagel defines a model for a theory *T* as a set of true propositions having the same formal structure or calculus as *T*.[6] May Brodbeck adopts a similar definition.[7] The position here is closely related to the previous claim that a model provides an interpretation for an uninterpreted formalism. According to the present claim this formalism is to be construed as that of some theory. The third assumption is that a model is in some sense at least to be construed as an analogy. Nagel, for example, uses the terms 'model' and 'analogy' interchangeably, and classifies the Bohr model of the atom and the wave model of light together with the analogy between molecules and billiard balls and the analogy between light and sound.[1] Braithwaite adopts a similar approach and cites models of molecules as linked systems of atoms and the analogy between hydrogen atoms and solar systems as examples of the same sort of thing.[2] Mary Hesse classifies the analogy between gas molecules and billiard balls together with postulates about atomic particles, claiming both as examples of models.[3]

---

[1] Ernest Nagel, *The Structure of Science*, New York, 1961, p. 90

[2] Ernest H. Hutten, *The Language of Modern Physics*, London, 1956, p. 82. See also his 'The Role of Models in Physics', this *Journal*, 4, 1954, pp. 284-301

[3] R. B. Braithwaite, *Scientific Explanation*, Cambridge, 1955, chap. 4; 'Models in the Empirical Sciences', in E. Nagel, P. Suppes, and A. Tarski (eds.), *Logic, Methodology, and Philosophy of Science*, Stanford, 1962, pp. 224-231

[4] 'Models in the Empirical Sciences', p. 225

[5] *Scientific Explanation*, p. 93

[6] Op. cit., p. 96

[7] 'Models, Meaning, and Theories', in L. Gross (ed.), *Symposium on Sociological Theory*, Evanston, 1959, p. 379

At the outset I must confess that I find all three of these assumptions unclear. A number of different interpretations might be placed on each, and it is not always apparent which is to be accepted. I shall discuss the first two of these assumptions in the present section, leaving the third for later.

To begin with some background is needed. The authors cited employ concepts of formal semantics in their analyses of scientific methodology. Following Carnap[4] they refer to a set of sentences obtained when a given calculus is interpreted as a *model* (or *interpretation*) for that calculus. They then apply Carnap's semantical distinctions to scientific theories, especially to those presented in an axiomatic manner. Here the axioms considered as uninterpreted strings of symbols can be distinguished from the same axioms considered as interpreted propositions. Using Carnap's terminology the former would be called a calculus, the latter a model. Now, if theoretical models of the sort we have been considering can be identified as models in the semantical sense, then the first claim above might appear warranted. For a model in the semantical sense does provide an interpretation for an otherwise uninterpreted calculus. If, in addition, the calculus in question is that of some theory then the second claim might also seem justified, for then the set of propositions comprising the model would have the same formal structure as that theory. But what reasons can be given for identifying theoretical models as semantical ones, what is involved in such an identification, and what is the warrant for holding that the

[1] Op. cit., chap. 6

[2] *Scientific Explanation*, p. 93

[3] 'Models in Physics', This *Journal*, 4, 1953, pp. 198-214; see pp. 199, 203. However, in her latest work, *Models and Analogies in Science*, London, 1963, two senses of the term 'model' are distinguished which appear to correspond to theoretical model and analogy, respectively; see chap. 1.

[4] Rudolf Carnap, *Introduction to Semantics*, Cambridge, 1942, p. 203

calculus alleged to be interpreted by a theoretical model is that of some theory? Here it may prove instructive to examine the discussion offered by Nagel, who perhaps more than any other representative of this position gives detailed attention to specific examples.

On p. 94 of the *Structure of Science*, Nagel writes:

> The Bohr theory of the atom . . . assumes that there are atoms, each of which is composed of a relatively heavy nucleus carrying a positive electric charge and a number of negatively charged electrons with smaller mass moving in approximately elliptic orbits with the nucleus at one of the foci. . . . The theory further assumes that there are only a discrete set of permissible orbits for the electrons, and that the diameters of the orbits are proportional to h^2n^2, where h is Planck's constant . . . and n is an integer. Moreover, the electromagnetic energy of an electron in an orbit depends on the diameter of the orbit. . . .

On p. 95, however, Nagel informs us that what he actually presented was not the Bohr *theory* but rather the Bohr *model* (' in such an exposition the postulates of the theory are embedded in a model or interpretation '). For the Bohr theory, we are informed, is really a calculus only some of whose terms are interpreted, whereas what is presented above is not a set of ' statement-forms ' but full-fledged ' statements at least part of whose content can be visually imagined '. (On p. 97, Nagel states that a model need not contain references to ' experimental concepts and observational procedures '.) Now we can certainly grant that what Nagel presents above is the Bohr *model* (as noted earlier, we can also grant that what he presents is the Bohr *theory*, as this expression is customarily used). Accordingly, given Nagel's notion of a theory it follows that not only does the Bohr model have the same calculus as the Bohr theory, but also that the Bohr model provides an interpretation for this (otherwise uninterpreted, or partially interpreted) calculus. I shall consider the latter claim first. This might be understood in a number of different ways.

(i) The Bohr model as an interpreted system can be distinguished from its calculus; it happens to be one among many possible interpretations that can be given for this calculus. If this is what is meant I have nó objection. A similar point can of course be made about any proposition or set of propositions. (' All metals melt ' can be distinguished from the schema ' all *A* is *B* ' and happens to be one among many possible interpretations that can be given for this schema.) More generally, I have no objection if Nagel and others simply want to distinguish the presentation of a theory (hypothesis, law) in an interpreted form from one in which (say) only mathematical equations with no meanings attached are given, where by a model they mean to be referring to a presentation of the former sort, nothing more. On the other hand, the examples cited by Nagel and others suggest that this is not merely their point. References to the Bohr model, the billiard ball model, the corpuscular model of light, models of molecules as linked systems of atoms, and so forth, suggest that they mean to be describing what is especially characteristic of those conceptions I have called theoretical models, not what is characteristic of *any* set of (interpreted) propositions whatever. Moreover, with respect to these models the suggestion appears to be not just that they are interpreted propositions distinguishable from their respective calculi, but something more. Nagel, we observed, says that the Bohr theory is a partially interpreted calculus, but that when ' customarily expounded ' it is ' embedded in a model '. This might be taken to suggest the following claim.

(ii) The *point* of formulating and utilising a model is to provide an interpretation for uninterpreted terms in a theory; or at least the model is actually used *because* it constitutes such an interpretation. Nagel recognises that this does not necessarily represent the historical order of development: he is not saying that scientists first formulate an uninterpreted or partially interpreted calculus and then supply a model. But he might be understood as saying that once a theory is conceived (no matter in what order) the model is actually utilised by the physicist to provide (or, because it provides) an interpretation for the calculus (of this theory). Such a claim, however, is too strong. The physicist no more utilises the Bohr model in order to supply an interpretation for a calculus (or because it supplies such an interpretation) than one who utters the sentence ' all metals melt ' does so in order to interpret the schema ' all *A* is *B* ' (or because the sentence provides such an interpretation). This is not to deny that the Bohr model *could* be used in certain cases to furnish (or because it furnishes) interpretations. One might, for example, consider certain mathematical equations and then show what physical interpretations they have in this model. But this context is a very special one, and what is envisaged is not something especially characteristic of theoretical models but holds equally for any set of propositions containing mathematical equations. Nor, again, is this to deny that certain models have played important roles in furnishing interpretations for formulas and may be used in certain instances because they furnish such interpretations. For example, the Schrödinger wave equation in quantum mechanics was interpreted in a particle model by specifying that the square of the psi-function at a given point in space represents the probability of finding the corresponding particle in the neighbourhood of that point. Three things should be noted, however.

First, this sort of interpretative role is not one characteristic only of models: the same role is often played by theories, many of which would not be classified as models. Thus, quantum statistics provides interpretations for certain terms and equations in thermodynamics. Second, when a theoretical model *is* used to provide (or because it provides) an interpretation for equations these equations will not be mere uninterpreted strings of symbols. The terms in the Schrödinger equation, for example, already have meanings though these may be fairly abstract; the particle model is then employed to assign a new interpretation to the psi-function which appears in this equation and is already interpreted as representing the amplitude of a matter wave at a point in space. Again, the Bohr model can be used to assign meaning to the term *R* appearing in Balmer's equation describing a relationship among the wavelengths of the lines in the visible region of the hydrogen spectrum. Yet this term (the gas constant) already has been given a meaning independently of the Bohr model by reference to the ideal gas law. Third, despite the fact that a theoretical model is sometimes used to supply (or because it supplies) new interpretations for terms in certain equations in the above ways, it would be misleading to consider this to be one of the principal reasons for its use, especially when considering those equations appearing in its own basic assumptions. The Bohr model, for example, is not something designed, or utilised by the physicist, to provide (or because it provides) interpretations new or otherwise for terms such as 'elliptical orbit', 'mass', or 'charge', which appear in its central postulates (indeed such terms were in use prior to this model and retain their original meanings in it). Rather it is a set of assumptions that employs these (already interpreted) terms to describe the atomic structure.

A final interpretation that might be placed on Nagel's claim is this.

(iii) Although scientists do not actually utilise partially interpreted calculi, nevertheless it is important for certain purposes to recognise the possibility of reconstructing the systems they do present in this way. If so, a model would be important in so far as it provides an interpretation for a system so reconstructed. In other words, if scientists were actually to present their ideas as partially interpreted calculi, then models would be utilised to serve the purpose specified. In assessing this claim the reasons for reconstructing scientific systems as partially interpreted calculi would need to be examined, a project beyond the scope of the present essay, though elsewhere I have offered criticisms of such an approach to scientific systems.[1] (It might be noted that for the very reasons generally offered in favour of reconstructing systems in the manner proposed—reasons having to do with the presence of ' theoretical ' terms—many theoretical models, such as atomic and nuclear models, would themselves need to be construed as partially interpreted calculi.) For present purposes, however, it suffices to note that even if it were valuable for scientists to reconstruct their ideas as partially interpreted calculi, and even if theoretical models were considered to be interpreted systems in the required sense, the interpretative role would not be one especially characteristic of such models, but would be one served by any set of scientific propositions. For any such set when presented as a partially interpreted calculus could then be utilised to interpret that calculus.

Turning now to the second assumption, viz. that a model has the same formal structure as that of some theory, various interpretations might be placed on Nagel's position:

(*a*) Consider any (axiomatised) set of propositions. Think of it as an uninterpreted calculus. Assign ' empirical ' meanings to *some* of the symbols involved and call this partially interpreted calculus a *theory*, since any partially interpreted calculus (whose interpreted terms refer to ' observables ') is to be called a theory. The original set of propositions will be a model for this theory. On this view, then, we obtain a model if and only if we interpret a calculus. (This interpretation might be suggested on the basis of Nagel's definition of a model (p. 96) and of a theory (pp. 90 ff.). It certainly appears to be the position of Braithwaite in *Models in the Empirical Sciences*, p. 231.)

(*b*) Consider any generally *recognised* scientific theory. Such a theory as ' customarily expounded ' is ' embedded in a model '. That is, what is usually presented is a set of (interpreted) propositions—the model. The theory itself is to be construed as the calculus for this set of propositions in which certain terms are interpreted (as designating ' observables '). From this point of view, we obtain a model if and only if we interpret the calculus of some generally recognised theory, where the interpretation will be one in terms of which that theory is usually presented. (This interpretation might be suggested on the basis of Nagel's discussion of the Bohr model.)

---

[1] ' Theoretical Terms and Partial Interpretation ', this *Journal*, 14 (1963), pp. 89-105; ' The Problem of Theoretical Terms ', *American Philosophical Quarterly*, July 1965

(*c*) Consider anything generally recognised as a scientific theory. We obtain a model if and only if we interpret the calculus of this theory, where such an interpretation may employ any meaningful notions whatever, and may indeed generate a description of a ' second system ' (p. 110) that is physically speaking completely dissimilar to the system described in the theory. (This interpretation, which includes (*b*) but not (*a*) as a special case, might be suggested on the basis of Nagel's later discussion on pp. 108 ff.)

Each of these claims implies that a model has the same formal structure as that of some theory. Yet none is acceptable if applied to what I have been calling theoretical models. For one thing none provides a *sufficient* condition. Consider any arbitrary calculus. Whether by interpreting this calculus a theoretical model is obtained (as claimed in (*a*)) depends upon whether or not the resulting propositions fulfil criteria for theoretical models expounded earlier. And this will not in general be assured, even if (as in (*b*) and (*c*)) the calculus in question is that of some recognised theory, and if, in addition (following (*b*)), the interpretation is one in terms of which that theory is usually presented. Furthermore, although (*a*) does provide a *necessary* condition for being a theoretical model it is a perfectly trivial one. Any theoretical model (or, for that matter, any set of propositions) is an interpretation of its own calculus. And if a theory can be constructed simply by re-interpreting some of the terms in this calculus, then any theoretical model (any set of propositions) has the same formal structure as a theory constructed by this process. Claims (*b*) and (*c*), on the other hand, are by no means trivial, since they require the theory in question to be one generally recognised as such and not one concocted simply by arbitrary re-interpretation of a calculus. Yet neither is acceptable as a necessary condition. For many theoretical models, for example, those of the atomic nucleus, are not interpretations of the calculus of any (known) theory of the sort required. To be sure one might wish to claim that the partially interpreted calculi of such models themselves are theories. But if so there would appear to be a violation of the requirement that the theory in question be one generally recognised as such; even more importantly, if this sort of partially interpreted calculus is to be classified as a theory, it is difficult to avoid the conclusion that conditions (*b*) and (*c*) would be trivially satisfied by virtually any set of ' empirical ' propositions.

3

In addition to theoretical models such as the Bohr model of the atom, Nagel cites the following examples in defence of his position: the laws of sound and those of the wave theory of light, the laws of heat conduction and those of gravitation, the laws of classical mechanics and those of relativistic mechanics, and other besides, where the first set of laws in each pair is what he refers to (interchangeably) as a model or analogy for the second.[1] By employing such examples in defence of his position Nagel appears to be claiming that the two major assumptions discussed in the preceding section hold for these cases as well. I shall consider this possibility first, and afterwards examine the third assumption referred to earlier, viz. that a model is in some way to be identified as an analogy.

To begin with, the claim, which follows from assumption 2, that the sets of laws cited above have the same calculus requires considerable modification. While some of the laws in each pair are formally similar others are not, or indeed may have no formal counterparts whether similar or dissimilar.[2] Also, even if their calculi were identical it would be misleading to claim, as in the case of the first assumption, that one set of laws provides an interpretation for the calculus of the other (unless this is understood in the trivial sense in which any set of propositions provides an interpretation for its own calculus). When, for example, a particle model is used to provide an interpretation (in more or less the ordinary sense) for the psi-function in the Schrödinger equation, it can then be said that the square of the psi-function in the Schrödinger theory represents, signifies, or means the probability of finding the corresponding particle. But we cannot in the same manner say, for example, that the expression ' centre of attraction ' in gravitational theory represents, signifies, or means a *source of heat* (though we can say that the former expression in the laws of gravitation corresponds to the latter in the laws of heat conduction). In short, one must distinguish between (i) showing that two terms in different sets of laws play formally similar roles, and (ii) providing an interpretation (representation, etc.) for a term in a given theory (as a semantical model is supposed to do for symbols in a calculus).

---

[1] Op. cit. pp. 108–114

[2] See my ' Models, Analogies, and Theories ', *Philosophy of Science*, **31**, January 1965.

What now of the assumption that a model is in some sense an analogy? To those accepting the characterisation of theoretical models offered at the beginning such a suggestion may seem puzzling. For it will be said, the Bohr model, for example, is something that attributes properties to a certain object or system, whereas the planetary analogy is essentially a comparison between such an object or system and a different one, the analogue. Moreover, the roles of each are different. The planetary analogy perhaps suggested some of the basic assumptions of the Bohr model; it is used in teaching the principles of this model, in helping one understand the atom as described by this model; it was helpful in suggesting how the model in its original form could be extended; none of which can be said of the Bohr model itself. The latter is used to explain certain spectral phenomena, to derive Balmer's equation, to predict additional spectral lines, etc., none of which can be said about the planetary (or any other) analogy. It might be objected, of course, that both provide *representations* of the atom. Even so, it is possible to distinguish between representing the atom by describing it as *being* a system satisfying Bohr's assumptions and representing the atom by describing it as *being like* a solar system in certain ways.[1]

What, then, might be involved in the proposed identification? Is it perhaps being suggested that a set of assumptions constitutes a model of *x* if and only if it is a description of an analogue for *x*? This would conform with the view suggested at one point by Nagel that a model describes, or can describe, a 'second system' distinct from that described in the corresponding theory. Yet if we consider many of the examples of models cited by Nagel and others the above is neither a necessary nor a sufficient condition. Many such models (of *x*) are not descriptions of analogues (for *x*). (The Bohr model of the atom describes atoms, not solar systems.) Nor is a description of an analogue necessarily a theoretical model. The laws of heat conduction describe phenomena analogous in certain respects to those in electrostatics. Yet these laws do not constitute a theoretical model in the sense in which the Bohr and billiard ball models do.

Perhaps, then, the identification of models and analogies is meant to suggest something weaker, viz. that the formulation of a theoretical model requires, among other things, that an analogy be invoked. Yet even this is dubious. No doubt an analogy will frequently be utilised and perhaps incorporated into the very formulation of a theoretical model, either explicitly or implicity (' electrons like planets revolve in elliptical orbits about an attracting object', 'the planetary electron revolves about the solar nucleus'). Use of simile and metaphor cannot fail to be noted and often proves helpful in learning and understanding the basic assumptions of the model. Still it is not something peculiar to theoretical models but holds for scientific hypotheses generally. Moreover, the assumptions constituting a theoretical model are frequently formulated without the aid of analogy, depending upon the level of the presentation.[1] If, then, the proposed identification is meant to suggest only that analogies *can* be important in formulating (and explaining) theoretical models, the point is well-founded (indeed, it was recognised in our earlier characterisation of such models). Stronger connections of the sorts mentioned above, however, appear to be neither necessary nor sufficient.

---

[1] Cf. Black, *Models and Metaphors*, p. 228

4

Having made these points about models and analogies some complicating factors must nevertheless be recognised. For one thing the terms 'model' and 'analogy' have frequently been used interchangeably by scientists and historians of science, especially in those cases in which an analogue of a given object or system is described in quantitative detail. For example, in his paper, 'On Faraday's Lines of Force', Clerk Maxwell describes an analogue of the electric field consisting of tubes of variable section through which flows an imaginary incompressible fluid satisfying equations analogous in certain respects to those governing electric fields.[2] The description of this analogue, developed, as Maxwell states, to establish theorems 'in a way more intelligible to many minds and more applicable to physical problems than that in which algebraic symbols alone are used', is sometimes referred to as a model of the electric field[3] (though Maxwell himself called it an analogy). Maxwell's assumptions concerning the imaginary fluid do resemble a theoretical model in the following respect: they attribute an underlying structure to the system they describe. Yet, by his own admission, this system is distinct from the electric field. Consequently the assumptions in question, though they may suggest various hypotheses about the electric field, cannot themselves be construed as a theoretical model of this field but only as a description of an analogue for it.

---

[1] See, for example, Robley D. Evans, *The Atomic Nucleus*, New York, 1955, for a presentation of nuclear models that is fairly free of the use of analogies.

[2] *Scientific Papers*, I, Cambridge, 1890, pp. 155-229

[3] See, for example, Sir Edmund Whittaker, *History of the Theories of Aether and Electricity*, I, New York, 1960, p. 279

A related point must also be emphasised. It is sometimes difficult to determine whether a given conception is being proposed as a theoretical model of a system x or as an analogy for x, i.e. whether what is being described is x or an analogue of x. This is important and may in part account for the proposed identification of models and analogies. Particularly good examples can be found if we turn from Maxwell's hydrodynamic analogue to later mechanical ' models ' of the ether.

The general aim here was to provide an entirely mechanical account of electric and magnetic forces. Two sorts of questions can be distinguished: (1) is a mechanical explanation possible? and (2) if so, which one ought to be accepted? At least some of the mechanical ' models ' of the ether must be understood as providing answers to the first question, not the second. Physicists formulating ' models ' of this sort were attempting to imagine mechanical systems that had the same effects as electromagnetic fields, while ignoring the plausibility of their particular conceptions (and even admitting their implausibility). For example, in his (later) paper, ' On Physical Lines of Force ', Maxwell introduces a number of assumptions concerning the ether, some of which I shall quote:

> (1) Magneto-electric phenomena are due to the existence of matter under certain conditions of motion or of pressure in every part of the magnetic field, and not to direct action at a distance between magnets or currents. . . .
> (2) The condition of any part of the field, through which lines of magnetic force pass, is one of unequal pressure in different directions. . . .
> (3) This inequality of pressure is produced by the existence in the medium of vortices or eddies. . . .
> (4) The vortices are separated from each other by a single layer of round particles. . . .
> (5) The particles forming the layer are in rolling contact with both the vortices which they separate. . . .
> (6) The effect of an electric current upon the surrounding medium is to make the vortices in contact with the current revolve.[1]

How does Maxwell treat these assumptions? Again I shall quote:

> My object in this paper is to clear the way for speculation . . . by investigating the mechanical results of certain states of tension and motion

---

[1] *Scientific Papers*, I, pp. 485-487

> in a medium [the ether], and comparing these with the observed phenomena of magnetism and electricity. By pointing out the mechanical consequences of such hypotheses, I hope to be of some use to those who consider the phenomena as due to the action of a medium, but are in doubt as to the relation of this hypothesis to the experimental laws already established, which have generally been expressed in the language of other hypotheses. . . . I do not bring it forward as a mode of connexion existing in nature, or even as that which I would willingly assent to as an electrical hypothesis. It is, however, a mode of connexion which is mechanically conceivable, and easily investigated, and it serves to bring out the actual mechanical connexions between the known electro-magnetic phenomena; so that I venture to say that any one who understands the provisional and temporary character of this hypothesis, will find himself rather helped than hindered by it in his search after the true interpretation of the phenomena.[1]

Has Maxwell formulated a theoretical model of the ether, or has he described simply a possible analogue for it? The answer is not clear-cut. If we examine his proposal in the light of features of theoretical models cited earlier we shall note similarities as well as differences. Like theoretical models Maxwell's conception consists of a set of assumptions about a system, viz. the ether. Moreover, these assumptions attribute an inner structure or mechanism to this system. Yet the spirit in which the conception is proposed, while somewhat akin to that of theoretical models, is different in important ways. For purposes of comparison we might formulate the gist of Maxwell's claim as follows: ' It is useful to represent the ether as having the structure I have described, for then various known principles can be derived from entirely mechanical assumptions; as to the actual structure of the ether I do not know, nor am I claiming it to be anything like what I have represented here; nevertheless, the fact that such a structure allows known principles governing electromagnetic fields to be derived shows the *possibility* of a mechanical theory.' The proponent of a theoretical model, on the other hand, is claiming that his representation provides at least some approximation to the actual structure of the system in question. This means that Maxwell's conception is in one respect at least akin to a description of an analogue. For by admitting that he does ' not bring [the proposed mechanical system] forward as a mode of connexion existing in nature ', he grants that this system is distinct from the ether. Still, unlike ordinary analogies, Maxwell is supposing for the moment at least that the mechanical system he describes actually constitutes electromagnetic fields, i.e. that they are not distinct but actually one and the same thing.[1] In an analogy such an identification is never made. While it may be supposed that an atom is *like* a solar system in certain respects the supposition is never made, even for the purpose of showing the possibility of a given type of theory, that the two are identical.

---

[1] *Scientific Papers*, I, pp. 452, 486

Although Maxwell's mechanical ether does not fit neatly into the classification of theoretical model of the ether or description of an ether-analogue, there are other nineteenth-century conceptions of the ether that do. In 1885, George Francis Fitzgerald formulated a ' model illustrating some properties of the ether '.[2] ' The model ', he writes, ' consists of a series of wheels rotating on axes fixed perpendicularly in a plane board, and connected by indiarubber bands '. FitzGerald made it clear that he did ' not intend it to be supposed that the ether is actually made up of wheels and indiarubber bands, nor even of paddle-wheels, with connecting canals '.[3] Moreover, unlike Maxwell in the example above, FitzGerald did not attribute the actual mechanism he described to the ether, even as a possibility. He was clearly constructing a mechanical analogue of the ether, and indeed often described what he was doing as drawing an analogy between the ether and the mechanical system he envisaged. In later years, however, FitzGerald proposed a different conception which he treated as a theoretical model of the ether:

> . . . I have been advocating the hypothesis that the ether is of the nature of a fluid full of vortical motion, and that electromagnetic actions are due to particular arrangements of this motion, which is, in general irregular or at least undirected. . . . What I now desire to call attention to is a hypothesis as to the nature of the wave motion which can be transmitted by a system of vortex filaments.[4]

The hypothesis in question, which attributes spiral vortices to the ether, was not proposed as a mere analogue, and unlike Maxwell, FitzGerald made no claim that he did ' not bring it forward as a mode of connexion existing in nature '. Instead he called it a ' hypothesis . . . put forward very tentatively '.

In sum, nineteenth-century speculations about the ether, despite the fact that they were called ' models ', varied considerably in conceptual import. Some were intended as theoretical models of the ether, some as descriptions of ether-analogues; still others had an intermediate status. One can look to features of the sort noted at the beginning of this paper in determining the status of a given proposal, even in recognising its borderline character. The existence of cases of the latter sort in no way vitiates the distinction between assumptions comprising a theoretical model of x and those describing an analogue of x, but shows only that such a distinction may be difficult to apply in certain instances.[1]

---

[1] This is to be contrasted with Maxwell's earlier proposal involving the imaginary incompressible fluid which is never assumed by him to comprise the ether but only to serve as an analogue for it.

[2] *Scientific Writings*, Dublin, 1902, p. 142

[3] Ibid. p. 151

[4] Ibid. pp. 472–473

The Johns Hopkins University,
Baltimore

---

[1] This paper was written as part of a project on scientific theories supported by the National Science Foundation.

MODELS AND THEORIES*

MARSHALL SPECTOR

1. Introduction

In this paper I will attempt to show that an important currently held view as to the nature of a model for a physical theory is in error. Indirectly, I shall be concerned to show that a certain related thesis about the nature of physical theories themselves is infected with serious difficulties. This will be done in the context of a modification of the former view which I shall offer.

It will be essential to have before us a careful statement of the latter thesis before we begin, for it seems to me that much of its apparent plausibility rests on the fact that presentations of it are not always very clear. (In fact, I believe that several influential attempts to *refute* this analysis of physical theories also rest on unclear and misleading presentations of it.)

According to this thesis, a physical theory is to be analysed as an empirically interpreted hypothetico-deductive system or formal calculus—in Rudolf Carnap's terms, a 'semantical system'.[1] Its most basic assumption is that a general distinction can be drawn between two types of terms occurring in physical theories—*observation* terms and *theoretical* terms.[2] The former, terms like 'green', 'desk',

---

* Received 21.ix.64.

[1] The oldest and most precise formulation of this position is found in Rudolf Carnap, 'Foundations of Logic and Mathematics', I, no. 3, of the *International Encyclopedia of Unified Science*, Chicago, 1939. Later statements of this view are found in other writings of Carnap; in various papers by Carl Hempel; in Ernest Nagel's *The Structure of Science*, New York, 1961; in R. B. Braithwaite's *Scientific Explanation*, Cambridge, 1953; Arthur Pap's *Introduction to the Philosophy of Science*, Glencoe, Ill., 1962; Ernest Hutten's *The Language of Modern Physics*, London, 1956; Peter Caws's *The Philosophy of Science*, Princeton, 1965. Further references may be found in the works just cited.

[2] See Carnap, loc. cit. p. 203; Carnap, 'The Methodological Character of Theoretical Concepts', in *Minnesota Studies in the Philosophy of Science*, I, p. 38; Nagel, op. cit. pp. 81 ff.; Braithwaite, op. cit. p. 51. The same assumption is made by the other authors mentioned in note 1. Most of them hasten to say that the distinction may not be a sharp one.

' longer than ', refer to observable objects, properties, relations, and events, and can be understood independently of any physical theory. The latter, expressions like ' electron ', ' magnetic field ', ' spin angular momentum ', refer to unobservable (= theoretical) objects, properties, etc., and can be understood only in the context of the theories in which they occur. The attractiveness of the view I shall be examining lies in its claim to present a general, well-articulated schema for showing just how we are to understand theories which seem to talk about objects which we have never observed, and perhaps never will directly observe. In outline, the analysis is as follows.

It is maintained that we can distinguish two components in any theory. The first is its *calculus,* which is the logical skeleton of the theory, considered as devoid of any empirical meaning. This will consist of a set of primitive formulas, i.e. sentences which are taken as postulates in the calculus; and other formulas which are obtained by derivation from the postulates in accordance with specified rules of transformation. Two types of terms appearing in the calculus may be distinguished: the primitive terms, i.e. terms which are not defined on the basis of other terms within the calculus; and non-primitive terms, i.e. those which are introduced on the basis of the primitives. (This distinction is not identical with that between observation terms and theoretical terms.) When this calculus, or syntactical system, is given an empirical interpretation, or meaning, it becomes a system of empirical statements having the structure of a hypothetico-deductive system. The primitive formulas become empirical hypotheses, and the derived formulas become empirical statements which will be true if the hypotheses are true.

To illustrate some of these ideas, consider the kinetic theory of gases. The postulates of this theory will contain such expressions as ' molecule ', ' mass of a molecule ', and ' position of a molecule '. These might be considered as primitives. Other expressions, such as ' momentum of a molecule ' and ' mean kinetic energy of a group of molecules ', will be introduced on the basis of the primitives. One of the postulates might be ' All gases are composed of molecules '. A typical derived formula might be ' If the pressure of a gas is increased while its temperature remains constant, its volume will decrease '.

The second component of a theory, the empirical interpretation, is given to the calculus by *semantical rules* for terms of the calculus, i.e. rules which are formulated in a suitable metalanguage (usually ' ordinary ' English, or German, etc.) and provides the meaning of the terms by stating what properties, relations, or individuals the terms designate.[1] An example is, ' The term " P " of the calculus designates the pressure of a sample of gas.'

Now the authors we shall be considering maintain that

(*a*) not all terms of the calculus of a theory *need* be given semantical rules, and

(*b*) not all terms of the calculus *can* be given semantical rules.

(It has not been clearly recognised that these are *distinct* claims. The importance of noting this distinction will become clear as we proceed.) Only the observation terms of the unanalysed theory are thus ' directly interpreted '. That is, if we look upon the calculus of a semantical system as the uninterpreted logical skeleton of a theory, for which the semantical system provides a *reconstruction*, only those terms in the calculus which represent the observation terms of the unreconstructed theory will be given semantical rules in the completed reconstruction. Theoretical terms of the unanalysed theory will not be given semantical rules in the semantical system reconstruction. It is claimed that such terms *cannot* be ' understood in themselves ', but must be understood—given their meaning—in an ' indirect ' manner, through the role they play in the theory. Such terms obtain an empirical meaning if and only if they appear in sentences in the calculus which also contain terms which *are* given semantical rules—the observation terms. Such sentences are known as *correspondence rules.*

An example of a correspondence rule is the postulate stated earlier: ' All gases are composed of molecules.' Symbolically, this would read ' $(x)(Gx \supset Qx)$ '. The term ' G ' is given the semantical rule ' " G " designates the property of being a sample of gas ', and the term ' Q ' (which is a symbolic translation of the theoretical expression ' is composed of molecules ') is not given a semantical rule, but is said to obtain a partial meaning indirectly by virtue of its occurrence in a sentence which contains a term (' G ') whose meaning is given directly and completely by a semantical rule. (It is important to notice that *all* the terms of a theory are found in the calculus, including the observation terms. If this is not kept in mind, there arises a tendency to confuse semantical rules with correspondence rules.)

This, in outline, is the analysis of physical theories which underlies the analysis of the concept of a model for a theory which I shall be examining. For reasons which should be apparent, I shall refer to it as the *partial interpretation thesis.*[1]

---

[1] See Carnap, ' Foundations of Logic and Mathematics ', loc. cit. p. 153 ; Carnap, *Meaning and Necessity*, Chicago, 1946, pp. 4 f. See also the introductory sections of Alonzo Church, *Introduction to Mathematical Logic*, Princeton, 1956.

The clearest and most precise explication of the concept of a model for a physical theory offered by a proponent of the partial interpretation thesis is that given by Braithwaite,[2] so the burden of my analysis will be directed towards his position. A model for a theory, according to Braithwaite, is to be understood as *another interpretation* of the theory's calculus, in which the *theoretical* terms are directly interpreted (by semantical rules).[3] It is sometimes also stated that these direct interpretations must be in terms of familiar concepts. But this need not be stated as a separate condition if we remember Braithwaite's (and Carnap's) injunction to the effect that this is the only way that direct interpretations can usefully be given.[4] Now if we use the term 'model' to designate not a domain of non-linguistic entities, but rather the statements about such entities, we can say that a theory and a model for the theory are two sets of statements which share the same calculus, but with the epistemological order of the two reversed. As Braithwaite puts it :

> A theory and a model for it . . . have the same formal structure, since theory and model are both represented by the same calculus. . . . But the theory and the model have different epistemological structures; in the model the logically prior premisses determine the meaning of the terms occurring in the representation in the calculus of the conclusions; in the theory the logically posterior consequences determine the meaning of the theoretical terms occurring in the representation in the calculus of the premisses.[5]

Braithwaite offers this explication of the concept of a model not as an arbitrary definition, but rather as

---

[1] I have borrowed this expression from Professor Peter Achinstein of the Johns Hopkins University.

[2] See his book, *Scientific Explanation*, as well as his more recent paper, 'Models in the Empirical Sciences', in Nagel, Suppes, and Tarski (eds.), *Logic, Methodology, and Philosophy of Science*, Stanford, 1960.

[3] It will not usually matter whether we use the term 'model' to refer to the domain of objects which the interpreted calculus makes statements about, or to the interpretative statements themselves.

[4] Cf. 'Foundations of Logic and Mathematics', p. 204. The metalanguage must be already understood. The familiarity of the concepts need not entail their observational character. They may be from another, 'better understood' theory ; but eventually, the chain must end with observational concepts. (See Braithwaite, 'Models in the Empirical Sciences', loc. cit. p. 227.)

[5] *Scientific Explanation*, p. 90

> an attempt to make more precise the notion of a model for a scientific theory widely current in discussions of the philosophy of science [ibid.].

Braithwaite also points out, quite emphatically, what models allegedly are *not*. Braithwaite claims that one of the chief ' dangers' in the use of models is the tendency that

> the theory will be identified with a model for it, so that the objects with which the model is concerned . . . will be supposed actually to be the same as the theoretical concepts of the theory [op. cit. p. 93].

Nagel also warns against confusing the model with the theory itself. After commenting on the possibility that a model may be ' an obstacle to the fruitful development of a theory ', he writes :

> The only point that can be affirmed with confidence is that a model for a theory is not the theory itself.[1]

Braithwaite goes on to say that

> Thinking of scientific theories by means of models is always *as-if* thinking; hydrogen atoms behave (in certain respects) as if they were solar systems each with an electronic planet revolving round a protonic sun. But hydrogen atoms are not solar systems; it is only useful to think of them as if they were such systems if one remembers all the time that they are not.[2]

According to the explication offered by Braithwaite, then, the objects of the model cannot be identified with the theoretical objects of the theory. Such an identification would be a logical error; the possibility of this identification (i.e. the question of the ' reality ' of the objects of the model) cannot even arise.[3]

However, as I shall argue, these questions *do* legitimately arise for some systems which physicists would recognise as models; and since Braithwaite's explication of the concept of a model cannot allow for this, it is inadequate. To show this, I shall begin by distinguishing four types of domains. They will be quite different in important respects; yet Braithwaite's explication will not be able to distinguish them. Then I will offer a modification of Braithwaite's explication which will be able to distinguish these types of domains. I shall conclude with an analysis of the effects of this modification on the partial interpretation thesis itself.

[1] *The Structure of Science*, p. 116

[2] *Scientific Explanation*, p. 93

[3] Compare Nagel, op. cit. p. 116

2. *Four 'Braithwaitean' Models*

(1) Suppose we have a geometrical interpretation of the calculus of some physical theory. (Braithwaite frequently uses geometrical interpretations of his 'factor-theories' as examples of models.) Such an interpretation would be a model for the theory in Braithwaite's sense, and yet the possible identification of the 'objects' of this system with the theoretical objects of the theory would not even arise for the physicist. There would be no question of the lines, triangles, circles, etc., being identified with, or being similar to, the theoretical objects of the theory. Only the formal structure would be relevant. If there are, in fact, triangles, etc., which satisfy the postulates of a given physical theory (its calculus), then the objects of this system are 'real'; but this does not mean that the theoretical objects of the theory are triangles, etc. This sort of identification would indeed be a mistake, but a strange sort of mistake that only a modern-day Pythagorean might make. (It is certainly quite different from the mistake of identifying the theoretical objects of the kinetic theory of gases with, say, billiard balls.) In this sense, a model has nothing whatsoever to do with the *domain* of the theory. Here, the system is another interpretation of the theory's calculus in a very strong sense of 'another'. The objects of the 'model' are of a different logical type from the objects of the domain of the theory. A physicist would probably not call this a model at all.

(2) It has been recognised that there is a rather thorough-going analogy between some of the laws of acoustical theory and those of electric circuit theory. Put another way, there is an important correspondence between acoustical systems and electric circuits. Let us consider a specific instance of this. If we have a series circuit consisting of a resistance R, a capacitance C, and an inductance L, with a periodic electromotance ϵ equal to $E \cos (wt)$, the charge on the capacitor, q, will satisfy the equation

$$\mathrm{L}\frac{d^2q}{dt^2} + R\frac{dq}{dt} + \frac{1}{C}q = E\cos(wt).$$

Now consider a Helmholtz resonator, which is a 'flask' of volume V with a neck of length d, radius a, and cross-sectional area $S = \pi a^2$. A sound wave of amplitude P impinges on the resonator opening, so that the driving force at the neck is given by $SP \cos (wt)$. The air displacement z at the neck will then satisfy the equation

$$(pd'S)\frac{d^2z}{dt^2}+\left(\frac{pck^2S^2}{2\pi}\right)\frac{dz}{dt}+\left(\frac{pc^2S^2}{V}\right)z = SP\cos(wt),$$

where $d' = d+\frac{16a}{2\pi}$, p = mean air density, c = sound wave velocity, and $k = w/c$. These two equations are of the same form:

$$\alpha\frac{d^2\xi}{dt^2}+\beta\frac{d\xi}{dt}+\gamma\xi = \delta\cos(wt).$$

For other acoustical systems, it is also possible to find electric circuit 'analogues'. In fact, this can be done in a general way, and certain combinations of acoustical parameters are given names taken from electric circuit theory. In the system just described, for example, the quantity $pck^2/2\pi$ is called the 'acoustical resistance', R; V/pc^2 is called the 'compliance', C; and pd'/S is called the 'inertance', M. Moreover, these names are not given merely on the basis of the expressions appearing in the same place in the acoustical equation as the corresponding expressions in the equation for the series circuit; the analogy is more than formal. Thus, the inductance of an electric circuit is a measure of the tendency of the circuit to resist changes in current ($= dq/dt$), while the inertance of the acoustical system is a measure of the tendency of the system to resist changes in air velocity ($= dz/dt$). Similarly, the capacitance of a circuit is a measure of the ability of charge to 'pile up', so to speak, in one place in the circuit, while the compliance of an acoustical system, e.g. the resonator, is a measure of just how far back the air in the neck will 'allow itself' to be pushed. (If it is claimed that the analogy is nevertheless a formal one only, this would strengthen the point I shall presently make.)

Finally, the analogy is actually used to solve practical problems. It is not merely a heuristic aid—allowing one theory to be taught in terms of the other. To quote from a textbook in acoustics:

> Consideration of the equivalent electric circuit offers many advantages in solving practical engineering problems of applied acoustics. For example, many acoustical systems are so complicated that their mathematical analysis is very difficult, if not impossible, and their design by a cut-and-try [sic] experimental method is extremely tedious, as each change involves constructing a new part. On the other hand, if it is possible to set up an equivalent electric circuit, the electrical constants of this network may readily be varied [actually—not just in the mathematical analysis] to obtain the desired experimental characteristics, and

> the constants of the mechanical system may then be calculated from their electrical equivalents. This technique has been used in the design of loudspeaker systems and other acoustical devices.[1]

Now, given all of this, is electric circuit theory a model for acoustical theory? (Or, are electric circuits models for acoustical theory?). Usually, the term ' analogue ' is used here, though one would be understood if the above relation were referred to as a modelling relation. The important point is, again, that there is no question of the identification of the subject matter of the two theories. Acoustics deals with small vibrations in air (generally, elastic fluids), whereas electric circuit theory deals with the movement of electric charge in certain types of systems. The analogy is perhaps not completely formal but the substantive similarities are rather weak.

(3) Here I have in mind systems which physicists would certainly recognise as models—systems for which the question of ' reality ' *does* arise. But here, the physicist knows that the objects of the model cannot be identified with the theoretical objects of the theory for *definite physical reasons*, although there is some *substantive similarity*. As an example of this, consider the following passage from the arch-modellist, William Thompson (Lord Kelvin). (I have italicised the parts to which I wish attention drawn.)

> To think of ponderable matter, imagine for a moment that we make a *rude* mechanical *model*. Let this be . . . [Here Kelvin describes a rather elaborate contraption, as Duhem might have called it] . . . you will have a *crude model*, as it were, of what Helmholtz makes the subject of his paper on anomalous dispersion. . . . If we had only dispersion to deal with there would be no difficulty in getting a full explanation by putting this *not* in a *rude* mechanical model form, but in a form which would commend itself to our judgment as presenting the *actual* mode of action of the particles, of gross matter, whatever they may be upon the luminiferous ether. . . . It seems that there must be something in this molecular hypothesis, and that as a mechanical symbol, it is certainly *not a mere* hypothesis, but a *reality*. But alas for the difficulties of the undulatory theory of light. . . .[2]

Kelvin has offered a model, which is only a ' rude ' model, *because* it will not work in certain important, fully specified, situations. If it were not for these other circumstances, which contradict the phenomena that could be expected on the basis of the model, this would not be a rude model, but a ' reality '. Consider also the following passage:

---

[1] L. E. Kinsler and A. R. Frey, *Fundamentals of Acoustics*, New York, 1950, p. 33

[2] *Baltimore Lectures on Molecular Dynamics and the Wave Theory of Light*, London, 1904, pp. 12–14

> The luminiferous ether we must imagine to be a substance which *so far as* luminiferous vibrations are concerned moves *as if* it were an elastic solid. That it moves *as if* it were an elastic solid *in respect to* the luminiferous vibrations, is the fundamental assumption of the wave theory of light [*loc. cit.* p. 9. (Italics mine)].

Kelvin does not identify his models with the domain of the relevant theory for *empirical* reasons—not because of a general philosophical point about the relation between models and theories. For Kelvin goes on to show that there are *other specific* phenomena with respect to which the ether does *not* behave like an elastic solid; (for example, the fact that material bodies move through it, showing that it sometimes behaves as if it were a fluid).[1] I think it is quite clear from the quoted passages that if it were not for these other, specific, intransigent phenomena, Kelvin would have dropped the as-if, rude-model, terminology.[2]

Notice that there is no question of identifying the ether with the actual gadgetry in a constructed model, just as for a 19th century physicist there is no question of identifying the molecules of the kinetic theory of gases with actual billiard balls (which are made of ivory, have numbers on them, etc.). The proposed identity would be between the molecules and elastic spheres, which are exemplified by billiard balls; or between the ether and an array of springs, etc., which is exemplified by the gadget on the laboratory bench.

---

[1] Thus it was much like the product once on the market in the United States known as ' Silly Putty '. This was a substance which could be moulded like clay, but which would also bounce if dropped several feet. It behaved like a fluid under some circumstances (slowly applied force), but like an elastic solid under others (rapidly applied force). Ice is another example ; glaciers flow, although ice is ' usually ' brittle.

[2] See also Edmund Whittaker, *A History of Theories of Aether and Electricity*, New York, 1960, I, ch. IX, ' Models of the Aether '. Whittaker describes a series of (mechanical) models of the ether, and it is apparent that the ' question of reality ' *did* arise for these models, and was taken quite seriously. In each of these cases, there was some *substantive* similarity between the domain of the model and the domain of the theory (electro-dynamics). And in each case, the model was considered only as an ' as-if ' because it didn't work—i.e. certain phenomena (though not all) expected on the basis of the model were not observed.

It is also interesting to note that in many of the models cited by Whittaker, part of the reason for failure was a lack of complete *formal* identity. Thus, according to Braithwaite's explication, certain paradigm cases of models employed by physicists would have to be described as *not* being models at all! But I shall let this pass.

Failure to make this distinction is what lends plausibility to Braithwaite's statement about hydrogen atoms and solar systems quoted near the end of the introduction to this paper. Of course hydrogen atoms are not solar systems; the nucleus of the hydrogen atom is not a star and the electron is not a planet—just as gas molecules are not billiard balls. But is it as obvious that the hydrogen atom is not a system in which something is going around something else? (Unfortunately there is no standard general term for such systems, analogous to the term ' elastic sphere ' to describe billiard balls and other such objects.) Bohr believed this to be the case; we no longer do for definite physical reasons, although the similarity is still there—hence the 'as-if'.

(4) Finally, there are systems which are no longer *merely* models. Here, the question of the reality of the model (i.e. the identification of the objects of the model with the theoretical objects of the theory) is not only significant, but is answered in the affirmative. A system may originally be introduced as a model in sense (3), but may eventually be modified to such an extent that we will finally speak of the identity of the objects of this system with the theoretical objects of the theory. When this happens, we may or may not continue to speak of the system as being a *model.* Actual usage may depend on factors such as the historical development of the theory. For example, we may still speak of the discrete particle model for the kinetic theory of gases, even though we have identified the concept of a molecule from the kinetic theory with the concept of a discrete particle (or ' object '). Here, the model is not *another* interpretation of the theory's calculus, but a *filling out of the original interpretation.* (This point will be made clearer in the following sections.)

Now, the first serious problem with Braithwaite's explication of the concept of a model as simply another interpretation of the theory's calculus is that it cannot distinguish between the four types of systems just discussed—and these *are* distinguished by physicists. For Braithwaite, all four would be models. However, we found that the first would not be recognised as such by physicists. The second would usually be described as an analogue. The fourth would probably not be described as a model either, but for a quite different reason—it is too good, as it were. ' Questions of reality ' arise for some of these systems, but not for others; whereas for Braithwaite, it is claimed that they should not arise for any of them. In other words, Braithwaite is correct in saying that identity of logical structure alone is not enough to be able to claim identity between the objects of the model and the theoretical objects of the theory. But I take this as showing that Braithwaite's explication of the concept of a model is inadequate, and not as an insight into the nature of models for physical theories.

It could be objected at this point that reliance on what physicists say in such situations is not sufficient for establishing a point about the logic of the situation (or for refuting such a point, as I am attempting to do). Perhaps not; I shall have more to say about this below. But suppose that we could emend Braithwaite's explication of the concept of a model in such a way as to take account of what physicists say—which *would* leave a place for the distinctions drawn above. And suppose that this emendation would be such that it could be accepted by the partial interpretation theorists as being within the letter and spirit of their analysis of the structure of physical theories. Clearly this would be a gain in understanding. I shall, in the next section, offer such a modification of Braithwaite's explication. I shall also attempt to show that this modification will be able to overcome certain further difficulties inherent in Braithwaite's explication.

3. *An Emendation of Braithwaite's Explication with an Example*

As long as only identity of formal structure is required between a theory and a model for it, there can be no relevant connection between the domain or subject matter of the theory and the domain of the model. A geometrical model for the calculus of a physical theory provides a striking example of this, and the electric circuit analogue for acoustical theory provides an example where the domains are both of physical objects, yet quite different. In each case, we do not have a model in the physicist's sense, and there is no question of the identification of the objects of the model with the theoretical objects of the theory.

If, however, the observable properties of the domain of the theory—the designata of the observation terms—are similar to the properties of the model represented by these same terms when the calculus is interpreted in the domain of the model, then the possibility arises of comparing the properties of the model represented by the theoretical terms of the calculus with the theoretical objects of the theory. That is, *we can argue by analogy to the nature of the theoretical properties.* A good example of this is provided by the elementary kinetic theory of gases and its usual model. It will be useful to compare in some detail the interpretations of the calculus of the kinetic theory of gases in the theory itself and in the elastic sphere model for the theory. (Notice that this distinction itself sounds somewhat strange, for this theory is usually presented by means of the model. The reasons for this will become clear as we proceed.)

In the theory, certain *defined* terms of the calculus are interpreted as designating the observable properties of a sample of gas in a container at equilibrium—volume, pressure, and temperature.[1] In the model, on the other hand, certain *primitive* terms of the calculus are interpreted as designating a group of elastic spheres in a container and some of their properties. Some of the primitive formulas express propositions describing these elastic spheres (their masses, for example), and others are interpreted as expressing the laws of classical dynamics (the laws governing their motion.)

Now we notice that the designata of some of the defined terms of the calculus, when used to represent the theory, are similar (in this case, identical) to the corresponding designata of these terms, when the calculus is used to represent statements about the model. Thus, for example, there is an expression in the calculus, which by the interpretation given the calculus in the model, represents the total rate of momentum transfer per unit area to the walls of the container in which the elastic spheres are moving. This expression, when the calculus is used to represent the theory, appears in a correspondence rule of biconditional form (a definition, in Carnap's sense) with the observation term 'P', which designates the pressure of the gas. But according to classical dynamics, rate of momentum transfer per unit area is equal to force per unit area, which is (by definition) the pressure on the wall of the container. Thus we not only have a formal identity—a shared calculus, but also a *substantive* identity of two properties, one from the domain of the theory and one from the domain of the model.[2]

The same is trivially true in the case of volume, thus giving us two sets of *identical properties*—out of a possible three.[3] We are left with temperature in the theory and mean kinetic energy of the elastic spheres in the model as the designata of a defined term (' T ') in the calculus when used in the theory and in the model, respectively. Are these last two properties also similar? Reasoning by analogy on the basis of the first two similarities (identities, in this case), we would expect so, which would amount to suspecting that gases are, in fact, *composed of* elastic spheres (suspecting that molecules are elastic spheres—that the model ' corresponds with reality '). At first sight, however, it would seem that the temperature of a gas and the mean kinetic energy of a swarm of elastic spheres are quite dissimilar. But we are not concerned here with temperature as a felt quality of bodies or as a sensation (hot-cold). Rather, we are concerned with the level of a column of mercury (for example—assuming that we are using a mercury thermometer) in a capillary tube. Now we notice that if liquids were also composed of elastic spheres,[1] they would expand if placed in contact with a gas, so composed, if (and only if) the mean kinetic energy of the particles of the gas were higher than in the liquid immediately before contact. In fact, an equilibrium state would soon be reached in which the mean kinetic energies of the two would be the same, at which time the liquid would cease expanding. All of this follows from classical dynamics. But this is just what is observed. When a mercury thermometer is inserted in a gas, the column will rise (or fall) for a time, and reach a stationary level.

---

[1] If pressure and temperature are not considered as sufficiently elementary or observable, the analysis could be carried out in terms of the observed heights of mercury columns ; but this would be at the expense of clarity without changing the results.

[2] Here, the identity is established on the basis of classical dynamics, rather than being an observed identity. But, as Braithwaite has noted, the familiarity need not necessarily involve observability. (See note 4 p. 124 supra.)

[3] We are at present interested in only three 'observable' properties of the theory—those which enter into the ideal gas law, $PV = kT$.

On the basis of this, and the two identities mentioned earlier, we can conclude at least that gases behave just as we would expect them to behave if they were in fact composed of small elastic spheres in incessant motion. That is, we can tentatively identify the objects of the model with the theoretical objects of the theory. Gases behave as if they were composed of elastic spheres. The ' as-if ' does not here indicate that we have contrary information (cf. Kelvin) but rather that we may still feel that there are other tests to which we would like to put this hypothesis before committing ourselves. If we had not the slightest idea of what other tests would be relevant, or if we could satisfy ourselves that other tests were impossible, there would be no point in the ' as-if '. In this case, there are such further tests, and they follow directly from a consideration of the elastic sphere model. (The theory itself—considered as the partially interpreted calculus with semantical rules for the observation terms only—affords no reason whatsoever to expect these other phenomena.) I have in mind the so-called 'molecular beam' experiments.

---

[1] But where the spheres also exert an attractive force on one another. This is admittedly crude, but putting in all of the details would only complicate matters without helping (or hindering) the point I wish to make.

We open a slit in one wall of the container of gas, and by a suitable experimental arrangement[1] we can see whether the results are those which would be expected (quantitatively as well as qualitatively) if a stream of small (unobserved) particles having a certain mean kinetic energy were to issue from the slit in accordance with the dynamics of such particles. As is well known, such experiments do in fact confirm the hypothesis that gases are composed as described—that we can identify the theoretical objects of the theory with the objects of the model. The measured velocities of the particles are what they would have to be in order to identify the temperature of the gas with the mean kinetic energies of these particles, thus completing the analogy.

Imagine a physicist who now says: 'Well, it has seemed as if gases might be composed of these small elastic spheres, considering the similarities you mentioned; now I'm convinced that they really are. After all, the results of this experiment too were exactly what could be expected if they were.' Has he committed an error? Has he succumbed to one of the 'dangers' involved in model thinking? If he had relied only on the *formal* characteristics of the model, he would indeed have made a mistake, although in this case a fortunate one. But the argument was based on the *substantive* characteristics of the model, and as such is perfectly good analogical reasoning. And it would still have been good reasoning even if the molecular beam experiment had given wholly different results, thus refuting the identification of the theoretical objects with the objects of the model. (Although we might then doubt the reliability of the apparatus rather than the model!) His mistake would then merely be that he had chosen a false hypothesis—but not a meaningless one, or an improbable one, or one which indicated a basic misunderstanding of the very use of models. The identification, on the basis of the first two sets of identical properties alone, is at least probable, and the results of the molecular beam experiment increase the probability.

Notice that it is not necessary at this point to produce a satisfactory theory of confirmation. Any explication of the notion of the probability of a scientific hypothesis which did not reflect the above would have to be rejected, as the reasoning sketched is as good a paradigm as can be found of one type of sound scientific reasoning.[1]

---

[1] See F. W. Sears, *Thermodynamics*, New York, 1955, p. 241—or almost any other elementary book on thermodynamics.

Thus, for a model to qualify as a ' candidate for reality ' (or, for a system to qualify as a model in the physicist's sense) it must be more than just some other interpretation of the theory's calculus. There must be a substantive similarity between the designata of the defined terms of the calculus when used to represent the observable properties in the domain of the theory and when used to represent the (formally) corresponding properties in the domain of the model. If there is, moreover, an identity of these properties, the model is no longer ' merely ' a model; the theoretical objects and properties simply *are*. . . .

This, then, is the emendation of Braithwaite's explication which will allow one to distinguish between the four types of domains sketched earlier, and which will accommodate an important type of reasoning based on models which physicists do in fact use. This emended form of Braithwaite's explication can also account for some of the uses to which models are put by physicists in modifying and extending a theory. I shall attempt to show this in the next section. But at the same time, it will become apparent that this emendation cannot be accepted by one who holds the partial interpretation thesis. We shall see that the modification which I have suggested is in contradiction with one of the most basic assumptions of this type of analysis of the structure of physical theories.

4. *Models and Modifications of a Theory*

The laws deduced from the simplified kinetic theory of gases sketched above are not as accurate as we would like them to be, and we would like to modify the theory to remedy this situation. In this case, the elastic sphere model ' points to its own extension ',[2] providing leads as to how it can reasonably be modified. For example, we might argue as follows:

---

[1] This type of reasoning is fruitfully compared with C. S. Peirce's ' abductive reasoning ' ; ' The surprising fact, C, is observed ; but if A were true, C would be a matter of course. Hence, there is reason to suspect A is true.' (*Collected Papers of C. S. Peirce*, Cambridge, Mass., 1935, 5, paragraph 189.)

Einstein's explanation of the Brownian motion on the basis of the kinetic theory of gases is a more striking example of this. It was this which convinced many doubters that the atomic theory was not just a ' convenient fiction '.

[2] This is Braithwaite's phrase in ' Models in the Empirical Sciences ', loc. cit. p. 229.

If gases are composed of elastic spheres, perhaps we should take their radii into account (i.e. they are not just point masses). Also, they may exert forces on each other when not in actual contact. If we make these assumptions, and modify the calculus of the theory accordingly, we will be able to derive a formula which, when interpreted term by term in the domain of the theory, is the van der Waals equation of state,

$$\left(P+\frac{a}{V^2}\right)(V-b)=cT,$$

where a is a constant for a particular type of gas depending on the force law, and b is a constant for each type of gas depending on the radii of the spheres. This is a more accurate description of the behaviour of gases than the original ideal gas law, $PV = kT$.

The new primitive formulas of the calculus (or, the primitive formulas of the new calculus) which allowed for the derivation of this law were essentially 'read off' from the model. By this I mean that a domain of objects was described which obeys the laws of another familiar theory—the dynamics of rigid spheres attracting each other in accordance with a stated force law (involving a). The statements of the description, together with the statements of the laws of this other theory, upon disinterpretation, become the primitive formulas of the modified calculus. (In the original model, the other theory was the dynamics of mass-points interacting only by contact. This was an idealisation, in an obvious sense, hence the name 'ideal gas law' for the equation of state derived from the associated theory. Note also how this is a reduction of part of thermodynamics to classical dynamics.)

In this modification of the model, the identity among derived properties in the model and in the theory, spoken of earlier, still obtains. Thus if we had started with the van der Waals equation of state, we could have used the same sort of reasoning by analogy described earlier to establish, or make probable, the identification of the primitive properties of this model with the theoretical properties of the theory. And once again, this sort of reasoning would not be valid if we had considered only the formal similarity (as Braithwaite correctly recognises when he argues against identifying the model with the theory). Moreover, if we had started with the van der Waals equation, it would have been impossible to *construct* the calculus of the kinetic theory without thinking of the model. (The very form in which the equation is stated betrays its origin in the model.) And if perchance someone hit upon the right modification of the original calculus without considering the substantive aspects of the model, he would have no justification for believing it probable.

This difference between what can be done with the ' theory itself ' and with the model is even more striking when we consider the partial interpretation thesis view that theoretical terms obtain their meaning by being connected with observation terms via correspondence rules. For now we have a much more complicated calculus just in order that the directly interpreted formula relating ' P ', ' V ', and ' T ' can be more accurate for a certain range of conditions. Now suppose (contrary to fact) that the ideal gas law were more accurate than the van der Waals law. Before dropping the modified theory in favour of the simpler one, physicists would, I think, look for experimental errors, because the modified model, associated with the modified theory, appears to be a *more plausible* representation of the theoretical objects than is the original model. But on the basis of the partial interpretation thesis, and its associated explication of the notion of a model, this would be an irrational procedure, stemming from a ' misunderstanding ' of what a model is.

Also, the derivation of the results of the molecular beam experiment could not be carried out with either the simplified or the full theory without a change in the observation language itself. We would need new terms in the calculus designating the distance between the container of gas and a sheet of film located inside a rotating cylinder, the angular velocity of the cylinder, the degree of blackening of the film, etc.[1] But if we introduce these new terms and the new correspondence rules needed for them, and at the same time accept the partial interpretation thesis as to how theoretical terms obtain their empirical meaning, we will have the very strange result that the meanings of the terms ' molecule ', ' mass ', ' velocity ', etc., have changed. After all, the theoretical terms are supposed to obtain their meaning from the observation terms through correspondence rules, and we have added new observation terms and correspondence rules. In the model, on the other hand, there is no such change in meaning. ' Elastic sphere ' means the same when the model is first conceived as it does after we realise that this model points to a new *test* of the theory in the molecular beam experiment. But we found that we could identify the objects of the model with the theoretical objects of the theory, through analogical reasoning based on substantive similarities in the two domains.

[1] See Sears, op. cit.

Therefore, on the basis of this identification we are forced to conclude (what is more plausible on its own merit) that the addition of the new test with its corresponding set of new observation terms and correspondence rules does not change the meaning of the theoretical terms of the theory, but rather gives new ways to test the truth of the theoretical statements (whose meanings are given in some other way—directly, as will soon become apparent.)

The view that the meanings of the theoretical terms change with the addition of new observation terms and correspondence rules has a similar strange consequence when we consider the sort of unification of physical theories that is accomplished by reasoning based on the use of a common model. For example, Niels Bohr, in his 1913 paper, 'On the Constitution of Atoms and Molecules,'[1] using reasoning based on a model of the atom (notice that it is not called a model for atomic theory) which can be considered as a detailing of the elastic sphere model of kinetic theory, explains why certain lines in the spectrum of hydrogen gas had been missing, and predicts under what conditions we can expect to observe them. According to the partial interpretation thesis, even though the Bohr theory contains the terms 'atom', 'mass', etc., these terms have a different meaning from that in the kinetic theory of gases. The observation language of the Bohr theory is different from that of the kinetic theory, as are (of course) the correspondence rules; and his reasoning, which depends on the identity of the meanings of these terms in the two theories, would be invalid.

In cases like this, it is the model which is the heart of the physicist's investigations. If we disinterpet the calculus that is read off from the model, and reinterpret it 'from the bottom up', and call the result the theory itself, looking now upon the model as merely another possible interpretation of the theory's calculus, we must accept the conclusion that arguments such as Bohr's are colossal logical blunders—fortunate blunders—showing a lack of understanding of how theoretical terms get their meaning, and a lack of understanding of what a physical theory is.

Now it may well be the case that much of what physicists *say* about the methodology of their science is of dubious value—I do not wish to argue this point here. But the partial interpretation thesis, coupled with Braithwaite's explication of the concept of a model (which leaves out substantive similarities), is not only in contradiction with this; it is also in contradiction with actual theory construction and theory modification—what physicists *do*. Now Braithwaite, Carnap, and Nagel say that they are not interested in constructing a logic of discovery. Rather, they are interested in analysing the final product. But their position is not neutral with respect to the former; it instead declares certain types of reasoning which are paradigms of physical genius to be (fortunate) logical mistakes of a very basic sort. (The parenthetical adjective alone should make one suspicious.)

---

[1] *Philosophical Magazine*, 26, 1913, pp. 9-10

We could not even describe the kinetic theory of gases as a reduction of thermodynamics to (statistical) dynamics. We could only notice that there is a formal similarity between some of the primitive formulas of the kinetic theory and the laws of classical dynamics. But these laws would not have the same meaning in each case. In the kinetic theory, the meanings of the terms ' force ', ' mass ', ' momentum ', etc., would have to be analysed in terms of the observation terms ' pressure ', ' temperature ', and ' volume '.

What has happened here? I think the basic point is that the suggested emendation of Braithwaite's explication of the notion of a model—i.e. taking into account substantive features of the model—as plausible as it may seem, cannot be accepted by one who holds the partial interpretation thesis. That is, even though my suggested emendation eliminates the difficulties I have been pointing out, and apparently does so while remaining within the letter and spirit of the programme of the partial interpretation theorists, the emendation contradicts a basic assumption of the partial interpretation thesis. For analogical reasoning from *substantive* similarities in the designata of terms in the derived formulas in the model and in the theory to substantive similarities in the theoretical (primitive) properties amounts to a direct interpretation of the theoretical terms in the theory. Thus, in the case of an *identity* of derived properties, the completion of the analogy is tantamount to *giving semantical rules for the theoretical terms.* (If there is only a similarity, in which case we have one type of ' as-if thinking ', we are giving qualified semantical rules, so to speak—cf. the earlier quotations from Kelvin.) But according to the partial interpretation thesis, we *cannot* give a consistent direct interpretation (semantical rules) for the theoretical terms of the calculus when it is used to express the theory, for ' we could not understand them ' (Carnap); the theoretical terms are ' not understood in themselves ', but only as part of the whole system (Braithwaite). Theoretical terms gain what empirical meaning they have *only* ' from below '; the 'zipper' moves from the bottom up in giving meaning to the theoretical terms. Thus, if this thesis is held as to how theoretical terms become meaningful, one is forced to consider a model as some *other* interpretation of the calculus. To accept our results is to accept the possibility of giving meaning directly for theoretical terms.

5. *Conclusions*

Our suggested emendation of Braithwaite's explication of the concept of a model has implied that theoretical terms *can* be given direct interpretations—semantical rules, where the metalanguage used to state them *is* 'understood'. It is in fact the same metalanguage used to interpret the observation terms of the calculus. That is, if we wish to analyse physical theories in terms of interpreted formal calculi, we *can* give semantical rules for theoretical terms (thus, for *all* terms of the calculus). Notice that this still gives the meanings of the theoretical terms 'on the basis of observations', but by statements outside the calculus rather than by correspondence rules.

I have also shown that theoretical terms *must* be given semantical rules, to allow for certain types of physical reasoning. Certain extensions and modifications of a theory were based on the use of a model *in the emended sense*, which we saw is tantamount to interpreting the theoretical terms of the theory. But this refutes what I referred to as assumptions (*a*) and (*b*) of the partial interpretation thesis in the introduction to this paper.[1]

All of this also implies that theoretical or unobservable *objects* may (in some cases) be described by observational *predicates*. Let us see why this is so. If a semantical rule is to be successful in giving meaning to a theoretical term of the calculus, it must be stated in a metalanguage which is 'already understood'. That is, it must supply a designatum which is understood independently of the theory being reconstructed. Otherwise it would be, as Carnap has put it, a useless transcription 'from a symbol in a symbolic calculus to a corresponding word

---

[1] I should point out here that in one sense the conclusions I have so far drawn, and those I shall presently draw, are not entirely new. (See, for example, R. Harré, *An Introduction to the Logic of the Sciences*, London, 1960, ch. 4 ; M. B. Hesse, *Models and Analogies in Science*, London, 1963 ; N. R. Campbell, *Physics, The Elements*, Cambridge, 1920, ch. 6 (recently reprinted in a paper-back edition by Dover entitled *Foundations of Science*) ; H. Putnam, 'What Theories Are Not', in *Proceedings of the Congress of Logic, Methodology, and Philosophy of Science*, Stanford, 1962, pp. 240-51.)

However, I believe that I have argued for these conclusions in an importantly expression in a calculus of words '.[1] This requirement has usually been put in the form of demanding that the properties, things, and relations which are to be the designata of the terms of the calculus must be observable, i.e. the metalanguage must be a theory-uninfected observation language. (This observation language is not to be confused with the observation language of the theory's calculus, which is part of the object language being analysed.) It would seem that familiarity should be sufficient; but let us grant the stronger requirement of observability for the present. Now, if we give semantical rules for the theoretical predicates (say) of the calculus, and do so in accordance with the just mentioned condition on the metalanguage, we will have the result that *unobservable objects may be characterised by observational predicates.* More accurately, we will have a sentence of the form '$P(a)$', with 'a' designating an unobservable (= theoretical) object and 'P' designating a property of which it is possible to observe instances applying to observable objects (e.g. 'having weight' as applied to an electron or atom).

This is the same conclusion reached earlier from a consideration of the concept of a model for a theory. We saw that the identification of the objects (and properties) of the model with the designata of the theoretical terms of 'the theory itself' was tantamount to giving semantical rules for theoretical terms and thus applying observational predicates to unobservable objects. In the example of the kinetic theory of gases, this amounted to maintaining that the unobservable atoms could have observational properties—or at least familiar properties (mass, velocity) from another theory (classical dynamics).

Moreover, we saw that there is nothing unintelligible about unobservable objects being characterised by observational predicates. The reasoning that leads a physicist to impute observational properties to unobservable objects was seen to be perfectly acceptable analogical reasoning.

These conclusions, however, conflict with the most basic presupposition of the proponents of the semantical system approach.

---

novel way. I have been concerned to show how these conclusions can be *generated out of* the 'partial interpretation' analysis of the structure of physical theories and models. I do not believe that this has been done by the authors cited above. (Hesse comes the closest to this sort of undertaking.) My *general* point of view, which may be called, roughly, 'realistic', is of course not new at all.

[1] Foundations of Logic and Mathematics ', p. 210

This is the assumption that we can distinguish in a general manner two types of terms or concepts (and statements) in physical theories: the observation terms, which designate observable properties, relations, events, and objects *only*; and the theoretical terms, which purport to designate unobservable properties, relations, events, and objects *only*. This 'dual language model' of the vocabulary of science has been expressed by Carnap as follows:

> [We accept] the customary and useful [division of] the language of science into two parts, the observation language and the theoretical language [, where] the observation language uses terms designating observable properties for the description of observable things and events [, and] the theoretical language [uses] terms which may refer to unobservable aspects or features of events, e.g. to micro-particles . . .[1]

We have seen, therefore, that what are perhaps the three most basic assumptions of the partial interpretation thesis are in error.

Now these assumptions involve two crucial notions—that of an observation term, and that of a theoretical term. In this paper, I have treated these notions as if they were clear and unproblematic. They are not, however, and a full evaluation of the partial interpretation thesis—and the semantical system approach itself—requires a careful analysis of them. This I hope to do in future papers.

[1] 'The Methodological Character of Theoretical Concepts', in *Minnesota Studies in the Philosophy of Science*, I, Minneapolis, 1956, p. 38

TOWARDS A THEORETICAL CONCEPTUALISATION OF SUPERSTITION *

KARL E. SCHEIBE and THEODORE R. SARBIN

" Man needs a rule for his will, and will invent
one if one be not given him "—William James

1. *Introduction*

THE purpose of this paper is to suggest a general way of thinking about superstitions. While it is common to describe some beliefs and behaviours as ' superstitions ', it is uncommon for philosophers, psychologists, and other systematic writers to be explicit about the meaning they intend to convey by the use of such a descriptive label. A wide range of phenomena, including magic, ritual, myth, and occult practices, is traditionally denoted by the term superstition, and by custom the whole lot is summarily assigned to the province of anthropological inquiry. While a great amount of concern within psychology has centred on such related topics as prejudice and attitudinal bias, not much effort has been directed towards an understanding of superstitions as such. Few and limited, indeed, have been the attempts to integrate a concept of superstition into psychological theory.[1]

To the theorist, the topic is immediately discouraging. The term ' superstition ' is so frequently used that common experience and common sense seem to provide an adequate basis for understanding. We know that superstitions are to be regarded as erroneous beliefs: irrational, primitive, and based on inadequate information. In addition, we subscribe to the assertion that superstitions, even though manifestly false, provide some want-satisfying, or utilitarian, function.

---

* Received 10.ii.65

[1] A notable exception is Willoughby (1935). A check of the Psychological Abstracts reveals that fewer than three articles per year have appeared under the ' belief ' index heading. Even fewer are listed under superstition. There seems to be no quantitative trend over the last thirty-six years. In a recent historical review of the topic, Cohen (1960) has shown that only five of a sample of fifty psychological textbooks have devoted a chapter to beliefs. The resistance of psychologists to investigate beliefs is worthy of study in its own right.

Then, too, we are aware that the amount of descriptive material existing on superstitions is vast by any standards. Frazer's (1925) encylopedic *Golden Bough* and Tylor's (1888) classic work on primitive societies are but two early examples of the bulk of ethnographic work on this topic. Lewinsohn's (1961) book on prophecy and prediction, Willoughby's (1935) lengthy article on magic, and Cohen's (1960) work on gambling behaviour are likewise replete with examples of superstitious practices. In short, the concept is too familiar to seem very strange. But upon examination this familiarity acquires a fascinating strangeness.

The focus of this fascination is the functional significance of invalid beliefs. A number of writers have been impressed with this idea. Malinowski's (1926) analysis of magical practices among the Trobriand Islanders centres about the thesis that every such practice and corresponding superstitious belief, ' fulfills some vital function, has some task to accomplish, represents an indispensable part within a working whole' (p. 132). Malinowski established the importance of superstitions by viewing them in their *functional* context. The serious student of culture and behaviour could no longer disdainfully dismiss superstitions as futile primitive beliefs or esoteric phenomena. Malinowski considered superstitions to be current vital elements within a culture rather than vestigial remnants of the past.

This conception is both tempered and extended toward psychology in Merton's discussion (1957) of the distinction between the manifest and latent functions of social behaviour. The application of latent function analysis clarifies apparent irrational, paradoxical, and superstitious social patterns. For instance, once the latent function of conspicuous consumption is seen to be the acquisition of status through expensive purchases, this pattern of social behaviour is better understood, and no longer appears paradoxical. Superstitions are established in this sense as continuous with normal or rational modes of believing and behaving.

A third and more purely psychological point of view is contained in Rotter's Social Learning Theory. Rotter's conception of ' Generalised Expectancies ' is inclusive of a class of beliefs that might well be labelled ' superstitions '. An expectancy is defined by Rotter (1954) as, ' The probability held by the individual that a particular reinforcement will occur as a function of a specific behavior on his part in a specific situation or situations ' (p. 107). Expectancies derived from specific situations tend to generalise to similar situations. Rotter views the failure to cross-validate generalised expectancies as responsible for superstitions.

The remaining root of the present discussion is a general inference model developed by Sarbin, Taft, and Bailey (1960) as a psychological description of the prediction process. Sarbin, *et al.* in considering the informational basis of categorical judgments (which are called instantiations), suggest that, 'Where adequate input is lacking, the cognitive organisations function to " fill in the gap " and instantiations are then largely a function of the modules (predicates) selected by factors other than sensory input' (p. 131). Propositions in the form of major premises stand ready to give meaning to sensory inputs. Propositions that are empirically derived or inductively confirmed earn the label ' scientific ' or ' rational '. Others, developed through analogy or fantasy or acquired through authority, until examined through induction, must be considered irrational or superstitious. Regardless of how derived, these premises serve as guides to action.

Both Rotter and Sarbin suggest, in effect, that a person's conduct is predicated on more information than is provided through current inputs. Malinowski and Merton argue that cultures and societies do the same thing. In all these sources, the thesis is emphasised that invalid beliefs have functional significance. We direct the rest of this essay to an elaboration of this thesis. In the end we hope to provide the basis for a useful set of conventions for the employment of the term ' superstition '.

2. *Some Definitions and Some Questions*

Anticipating the model which will later be presented, we define a superstition as a special kind of expectation (or subjective probability, hypothesis, premise, belief). A superstition may be said to exist whenever an individual persistently or repeatedly *behaves as if his subjective estimate of the result of that behaviour is significantly different from an objective (scientific) estimate of the effect of that behaviour.* Some unsettled problems as to the inclusiveness of this definition are contained here, but it will do for a start. However, it should be clearly understood that the term superstition is not meant at this time in its usual pejorative sense. The full reasons for this will become clear later.

Questions that deserve examination regarding these special expectancies include the following. What determines whether or not an individual subscribes to a given superstition? Some people carry charms and amulets, consult astrologers and palmists, and some do not. Is it the case that some people are superstitious and some are not? Or may it be that everyone is superstitious in nearly like degree, only the specific content of the superstition varying from individual to individual and from group to group? That college students subscribe to numerous superstitions had been demonstrated (Gilliland, 1930; Dockeray and Valentine, 1935). Is it not appropriate to describe the postulates of scientific theoreticians and of theologians as somewhat highly evolved superstitions?[1]

At the moment, we have no systematic studies to support hypotheses regarding differential distributions of superstitions among different educational classes.

These problems are approachable, but it seems that they may all be subsumed under the more general question: What are the conditions under which superstitions are formed, maintained, and abandoned? It is to this question that we address ourselves, but not until we present a theoretical conceptualisation of superstition in more detail.

3. *A Cognitive Model*

What does a man need in order to participate in a culture? In the simplest terms, he needs a body with muscles that work. In addition, he needs some desires, or values. These two propositions may be combined into the following sentence: Man wants a great variety of things; for example, food, shelter, and clothing; children, happiness, satisfaction, security, and a sense of self-esteem, and is capable of action. But he lacks the *sine qua non* to participate in a culture, to behave effectively. He needs *to know* how to direct his actions toward desired ends. In Tolman's (1958) terms, he needs to have a cognitive map of means-ends relationships. In the terms of the present analysis, he needs a set of beliefs, some of which may be denoted by the term 'superstitions'.

The behaviour model here suggested is in part borrowed from Rotter (1954). The *potential to behave* is some function of the expectancy regarding a favourable outcome of that behaviour, and the *value* of the outcome of that behaviour. Expectancies derive from past observations of contiguously occurring events, and may be said to represent *perceived* cause and effect relationships between behaviours and the outcomes of behaviours. To complete the skeleton of the model for superstition, it need only be observed that *real* cause and effect relationships are obscure, and often quite different from perceived cause and effect relationships.

---

[1] Since the pejorative sense is not here intended, an affirmative answer to this last proposition need offend no one.

We are now brought up to the *psychological* problem of epistomology. How do we come to know *what* we know? What is the relation of the real world to the known world? How do we know *that* we know? Hume argues that cause and effect relationships are not ultimately compelled upon us *in toto*, but rather are inferences based upon contiguity. He says in his *Treatise of Human Nature*, ' We find only that the effect does, in fact, follow the cause. The impact of one billiard ball upon another is followed by the motion of the second. There is here contiguity in space and time, but nothing to suggest necessary connection.' Yet, it is often over-looked that though this inference be logically unnecessary, it is psychologically very powerful. Cause and effect relationships *are* psychologically compelled upon us *in toto*, as Heider (1944), Michotte (1963), and Piaget (1955) have amply demonstrated. In short, it is possible to believe that the known world is the real world, and to act as if it is, even though it might be impossible to be ultimately correct in this belief.

Consider an example: Farmer Brown wants to plant his corn. He wants to do it so as to maximise his autumn harvest. What is to determine when and how he plants and cultivates his corn? He certainly does not proceed on a trial and error basis. He uses a number of rules. These will have to do with the calendar date, the state of the weather, the moisture of the soil, the type of seed, and perhaps the phase of the moon or the pain in his bunions. He has probably derived few if any of these rules by his own empirical observations, but has acquired them from his ' cultures '. But only some of them would be called superstitions, and these would be so called because of some demonstration that the application of the rule has no effect on the outcome of the crop. However, even if Farmer Brown were to be shown compelling statistics which contradicted one of his rules (expectancies) he still might apply the rule. Under these circumstances, we would feel justified in calling Farmer Brown's behaviour superstitious.

To rephrase the conceptual scheme: superstitions come to exist because of imperfect knowledge of antecedent-consequent relationships in the ecology. On the assumption that an elaborate and fairly definite set of beliefs is *necessary* to support active behaviour, we consider superstitions to ' fill the gap ', as it were, between the objective probability estimates available to a person, and the degree of subjective probability required to execute behaviours. The beginning fact is objective uncertainty. In the face of this uncertainty, beliefs must be adopted, so that a man is not transfixed for lack of Truth. Where he lacks truth and needs it, he makes it up, somehow. We shall now consider how this construction occurs.

4. *The Development of Superstitions*

In an experiment with pigeons, Skinner (1948; also Morse and Skinner, 1957) has demonstrated the means of development of a ' sort of superstition '. Using his operant conditioning technique, he was able to establish firmly those behaviour patterns which *happened* to precede the presentation of the reinforcement. The bird behaves as if ' there were a causal relation between its behaviour and the presentation of food, although such a relation is lacking ' (Skinner, 1948, p. 170). The superstitious rituals are due not only ' to the fact that a reinforcing stimulus strengthens any behavior it may happen to follow, even though a contingency has not been explicitly arranged, but also to the fact that the change in behavior resulting from one accidental contingency makes similar accidents more probable ' (Morse and Skinner, 1957, p. 308).

Harry Stack Sullivan quotes as a prime example of primitive ' parataxic ' thinking a case portrayed in one of Franz Kafka's short stories. A dog who lived in a kennel surrounded by a high fence was urinating one day when a bone was thrown over the fence. The dog thought, ' My urinating made that bone appear '. Thereafter whenever he wanted something to eat he lifted his leg.[1] Sullivan believes that much of our thinking does not advance beyond the level of parataxis; that we see causal connections between experiences that have nothing to do with one another.

The essential features of these animal illustrations is that accidental contingencies are mis-represented as essential contingencies. There is, of course, a wealth of anecdotal human evidence which fits the same description. For example (from Diprose's *Anecdotes of Curious Superstititions and Omens*):

> In Devonshire, the appearance of a white-breasted bird has long been considered an omen of death. This belief has been traced to a circumstance said to have happened to the Oxenham family in that county, and related by Howell, in his Familiar Letters; wherein is the following monumental inscription : Here lies John Oxenham, a goodly young

---

[1] We are indebted to Hall and Lindzey (1957) for this example.

> man, in whose chambers, as he was struggling with the pangs of death, a bird with a white breast, was seen fluttering about his bed, and so vanished. He died immediately. The same circumstances are related of his sister Mary, and two or three others of the family.

The paradigm is the same. First is the condition of uncertainty regarding a valued outcome: Skinner's pigeons uncertain as to when or whether they were going to be fed, and John Oxenham's relatives uncertain as to when or whether John would expire. Then, immediately prior to the revelation of the outcome, something fortuitous happened. Moreover, this fortuitous pairing of events was repeated, at least once. Since both the pigeon and the Oxenhams are ignorant of any more powerful causal agent, the fortuitous events are taken as causes. In the future, the pigeon has the habit of making swiping motions with its head. The Oxenhams have the habit of avoiding white-breasted birds. The same is true of Kafka's dog, which has the habit of raising its leg at inappropriate times.

This type of explanation does well enough for the individuals concerned. But how do we explain the fact that the white-bird superstition becomes widely disseminated throughout the culture, while the colony of pigeons is probably not so taken in? We would do well at this point to heed a warning made long ago by Jastrow (1900) regarding the development of superstitions. He warns against the practice of analogical reasoning: The inference of a further degree of resemblance from an observed degree of resemblance. Jastrow argued that many superstitions, in particular Voodoo practices, derive from this faulty logical procedure. Sarbin, *et al.* (1960), have formalised this process of analogical mediation as follows:

> A has characteristic x.
> x and y are defining characteristics of species M.
> Therefore, A has characteristic y and is a member of species M (p. 62).

The conclusion, of course, is false. Yet the Voodoo Priest persists in the belief that mutilation of the representation of a person will do harm to that person, and our more primitive comparative psychologists persist in the belief that what is true of learning in the pigeon is also true of human learning.

We should conclude that the pigeon can carry us no farther in our understanding of superstition. We are left with the majority of human superstitions quite unexplained. Skinnerian conditioning must involve the empirical contiguity of the perceived cause and effect for each individual who bears the superstition. Fortunately or unfortunately, there seems to be something else than this involved in the ways people acquire superstitions. In fact, for human beings, even the first-order demonstration of the efficacy of superstition is not necessary. That is, they need not ever be witnesses to the original contiguity between the fortuitous event and the outcome. For example, Leonardo da Vinci, in *The Last Supper*, painted Judas Iscariot in the act of knocking over the salt-cellar. This is held to be the origin of the still extant superstition that it is bad luck to spill the salt. Now most people who have this superstition are probably not even aware of its origin. Likewise, astrological superstitions may have had their origin in the fortuitous pairing of a given stellar orientation and a particular noteworthy earthly event. But such origins are buried in antiquity. Yet people continue to be superstitious about the placement of stars.

It is to be noted that popular astrology has enjoyed a recent boom and revival. Lewinsohn (1961) offers the following speculation on this phenomenon:

> While the revolution which has been changing the Newtonian world picture ever since 1900 has made some people more skeptical, it has simply made others more credulous. Scientific laws which our textbooks have taught us to look upon as eternal truths have collapsed like ninepins, and new ones still lack the authority that their predecessors enjoyed. Scientists have grown a little more circumspect and no longer dare to proclaim general laws with the facility of their 19th century predecessors. Because of these developments, laymen have begun to feel that it is impossible to distinguish truth from hypothesis. Now, laymen have always been sticklers for certainty, and where truth is uncertain, they will accept anything that is presented as being simple and sure (pp. 93-94).

In this vein we might also suggest that the current resurgence of mysticism in psychology, as exemplified by some of the existential psychologies, and the popular issue of consciousness-expanding, or psychodelic drugs, might be traceable to the lack of definitive statements within the field of psychology, to the disillusionment with psychoanalytic absolutes, and to the great difficulty of understanding current psychological orientations.

Returning to the problem of the acquisition of superstitions, we shall no longer avoid the most simple declaration: that is, people may acquire their superstitions by direct verbal communication, as well as by first-hand perception of fortuitous pairings of events. It is as if each culture presents its participants with a rule book which states standard beliefs about what leads to what. These standard beliefs may be regarded as major premises that stand ready—under conditions of uncertainty—to instantiate a potentially threatening or disturbing event in such a way as to bring it under control.

5. *The Maintenance of Superstitions*

> Any object which remains uncontradicted is *ipso facto* believed and posited as absolute reality—William James

It is the thesis of this section that if an individual already has an explanation regarding a given event, that explanation will tend to stand until replaced by some ' stronger magic ', and that this replacement is contingent on several conditions. It is commonplace in science that a theory is rarely, if ever, negated by its own terms, but rather must be replaced by a better theory if it is to be replaced at all. Perhaps this is why Freudian theory has endured. Volumes of criticism have been directed towards it, but nobody seems to have publicised a competing theory as general and elaborate and dramatic for the understanding of psychological cause-effect relationships. Perhaps some of the newer psychodynamic theories are just as general, but still the Freudian monument stands, and will continue to stand through its own inertia, unless it is challenged by a significantly stronger (more dramatic, more need-fulfilling) theory. Little is gained in substituting one belief for a pragmatically equivalent one.

One might point to a set of beliefs held by laymen and by many professionals in the mental health field as an example of the maintenance of a superstition. Some disordered persons (i.e. those whose behaviour does not follow rules considered appropriate by certain members of the society) are ' diagnosed ' as sick persons and are given a label—schizophrenia. This labelling is coterminous with what goes on when a person is sick as the result of microbial invasion, virus infection, fractured bones, etc. Because of superficial similarities in the conduct of some disordered persons—disturbed speech, invalid judgments, ' hallucination ', etc.—with the conduct of persons with known disease processes (toxins, brain tumours, virus infections, etc.), the illicit analogical leap is made that all disorder may be accounted for through application of the disease model. The failure to establish uniform diagnostic criteria for schizophrenia, the failure to find language and thought disorders substantially different from rates in the normal population (when the effects of commitment and hospitalisation procedures are controlled), and the arbitrariness with which a person is classified as mentally ill, have not eliminated the belief that disordered persons are ' sick ', must be ' diagnosed ', and subjected to legal and medical procedures—presumably for the ' cure ' of the patient (Sarbin, 1962). The operation of such a superstition to ' fill the gap ' is epitomized in Cumming and Cumming's (1955) account of the ineffectiveness of a community mental-health educational programme.

> In our attempts to produce a more permissive climate for former mental patients, we conveyed the idea that they were pretty much like everyone else, and that there was no sharp line dividing the sane from the insane, but rather a continuous range of behavior. In stressing this idea we were hammering directly at the core of the community's own solution to the problem of the mentally ill. Our problem was not theirs. We were concerned with the cure of the mentally ill, the people of Prairie Town with the stability and solidarity of their own community. . . . Prairie Town's pattern of beliefs and attitudes toward mental illness was not merely a patchwork of half-truths, fallacies, and inconsistencies. . . . it played an important part in preserving the well-being of the community and the peace of mind and self-esteem of the average individual (pp. 63-64).

The basis for the longevity of superstitions need no longer be shrouded in mystery. If we accept the general statement of the Law of Effect—a response that is followed by reinforcements will probably be employed again under the same or similar circumstances—then the events that support superstitions may be ordered through the inference model (Sarbin, *et al.* 1960). Certain events occur. They provide sensory inputs. If these inputs cannot be readily instantiated (assigned to a class) by the participants in a culture, cognitive strain follows. This is characterised by large increases in cognitive activity—an effort to find a major premise that contains the same middle term as the so-far-uninstantiated input. Increases in cognitive activity frequently ' spill over ' into physiological activity, sometimes identified as dysphoric affects. The solution to the problem is through the employment of ritual or other procedures dictated by the superstition. Filling the gap in knowledge through provision of a usable major premise reduces cognitive strain. This may be reinforcing in either or both of two ways (1) it may be arousal reducing (drive reduction in the Hullian sense) and (2) the persons carrying out the ritual are given approval (generalised social reinforcement in the Skinnerian sense). The solution—no matter how irrational from our point of view—is thus self-perpetuating. Following are some instances of this general theme.

Among the Polynesians it is very common to resolve the questions of guilt or innocence of a defendant by spinning a coconut, and observing the direction in which it comes to rest. It is common knowledge that witch trials were often conducted by the similar method of 'ordeal', in which the accused was thrown into water, or otherwise placed in a perilous situation. Survival was accepted as proof of innocence, while death gave certainty of guilt. After the ordeal is over, the accepted conclusion is a source of satisfaction to all concerned. If the accused survives, the accusers are relieved that they have not taken a life in vain. If the accused dies, the actions of the accusers are vindicated, and again relief, or a reinforcing state of affairs, is the consequence.

Or consider the example of prayer to believed-in gods. Since the prayer is believed to be efficacious in altering the course of future events, the worshipper feels relieved and comforted immediately after having prayed. He may now rest because he has placed control in the hands of a greater power. An appeal has been made to the powers which control all uncertainty, and human worry is no longer necessary. Even if the prayer is not answered the prayer is still, on the whole, a source of relief and comfort. Further, his actions may be approved by significant others in the community.

Or consider the more homely and current example of filter-tip cigarettes. It is now fairly well established that the presence of a filter does little to alter the probability of smoking leading to lung cancer (*Smoking and Health,* 1964). Yet filter cigarettes succeed in reducing some of the tension in smokers aroused by the cancer research findings.

Additional examples could be cited by the score, but the essential point is that the superstition does indeed have an effect, even though the perceived causal relation is entirely false. Once installed, and invested with a high degree of confidence, the superstition is completely self-maintaining, and requires no further confirmation in the way of fortuitous pairings of events, so long as the superstition is not subjected to conflict and challenged by an alternative and functionally more appropriate belief.

The Hutterite communities of the eastern United States provide a revealing example on this point[1]: the culture remains quite distinct from that of the surrounding communities, and this distinction is cherished by the Hutterites. A remarkable people, they have successfully resisted the intrusion of science into the majority of their customs, and do not seem much the worse for this rigidity. In fact, their childbirth traditions function so well for them, that while the average mother has 11 children, a severe complication is very rare. The Hutterites have numerous superstitions about the ' evil eye ' and about sexual taboos, and these are accepted from generation to generation without much question. Alcoholism, crime, dope-addiction, and mental disorders are present only to a negligible degree. In fact, the current problems of this culture can be traced directly to the partially successful attempts of outsiders to have them institute more modern practices in childbirth and child rearing. The beliefs of this culture have worked for them in a very efficient and satisfying fashion and are thereby self-sustaining. The truth of this statement is demonstrable even though we would not hesitate to label a large number of their beliefs as ' superstitions '.

---

[1] The source of information here is a report on 'folk obstetrics' by Eaton (1958).

Of course, not all persistent superstitions have such an obvious functional significance. Saunders (1958) reports that beliefs regarding medical practices came from four widely disparate sources for the Spanish-Americans of the south-west. These people seem to be able to accept all of these beliefs, even though many of them seem contradictory. Characteristically, they try their own folk-medicine first, and if these techniques are ineffective, they simply turn to something else. ' A remedy is tried, and if it works no surprise is evinced, since that is what was expected. If it does not work, the failure is rationalized and something else is tried ' (Saunders, 1958, p. 191). Clearly these folk beliefs exhibit a fair amount of inertia in the face of challenge. Beliefs which seem inconsistent and dissonant to us are not necessarily so for the people who hold those beliefs.

Another basis for the maintenance of a superstition is the process of insulation against or distortion of objective information to conform to the superstitious commitment. For example, for some individuals agreement between the conclusion of a syllogism and one's own beliefs functions as a criterion for judging the soundness of syllogistic reasoning (Thouless, 1959; Morgan and Morton, 1944; Thistlethwaite, 1950). Highly prejudiced individuals are resistant to change in general (Rokeach, 1943) and individuals with highly conservative beliefs are resistant to contrary information (Frenkel-Brunswik, 1949).

It has been experimentally demonstrated that the extent to which beliefs about matters of fact deviate from the truth is highly related to the level of information the respondents possess about those facts. Moreover, the distortions in beliefs are consistent with the evaluative significance assigned by the respondents to those beliefs. That is, individuals tend to believe what they want to believe where lack of information makes this possible (Scheibe, 1963).

In sum, superstitious beliefs seem to have a fair amount of stability. If they are not challenged they are maintained. Even if moderate degrees of challenge are made to established superstitions, the individual is likely to prefer distorting facts or logic to discarding the superstition (Cohen, Brehm and Latane, 1959). Of course, there are interesting individual differences in this regard: 'The passion for unity and smoothness is in some minds so insatiate that, in spite of the logical clearness of... reasonings and conclusions, many will fail to be influenced by them' (James, I, 1890, p. 162).

6. *Another Function of Superstitions*

A tertiary functional significance of uncertainty may give rise to superstition. James has said that uncertainty must exist in order for man to have a sense of personal significance. Many 'classical' psychologists would agree with him. (For instance, France, 1902). The essence of Kierkegaard's philosophy is that man must cope with uncertainty and unpredictability. If man were to have true vision of cause and effect relationships, there would follow, according to some, a complete abrogation of human responsibility. It is the uncertainties which give life its vigour and interest. Others of a less romantic persuasion view this line of thinking as a severe restriction on the progress of scientific knowledge, and openly propose that we seek to establish a completely controlled and predictable universe. The flavour of this issue is conveyed in this statement by Skinner (Rogers and Skinner, 1956).

> A world in which people are wise and good, without trying, without 'having to be' or without 'choosing to be', could conceivably be a far better world for everyone. In such a world we should not have to 'give anyone credit'—we should not need to admire anyone—for being wise and good. From our present point of view we cannot believe that such a world would be admirable. We do not permit ourselves to imagine what it would be like (p. 1060).

We do not propose to discuss this issue here. However, it should serve as a caution to those who would cavalierly declare that human beings should be relieved of their uncertainties and of the superstitions which are their product.

To recapitulate:

First, a superstition may be considered to be a special kind of expectancy that can be codified as a major premise. Starting with the assumption of objective uncertainty, it has been argued that some extra-logical propositions, such as superstitions, are necessary guides to action.

Though initially superstitions may be *acquired* in a manner analogous to the process of operant conditioning, such a model does not do well in describing most superstitions, especially those for which the original confirmation is historically obscure. We have suggested the possibility that superstitions may be acquired by direct tuition on condition of uncertainty of outcomes, where the outcomes are highly valued.

If a superstition goes unchallenged by another and ' stronger ' superstition, it will not just wither away. Superstitions do have the effect of reducing cognitive strain and thus are self-maintaining. They also exhibit a high degree of inertia, in that once they are installed, they then remain until dislodged by a much stronger belief. Individuals and cultures may differ in the extent to which they can tolerate the ambiguity of contradictory superstitions.

It should be clear that the necessary conditions for the application of the superstition label are located *both* in the actor and in the observer. If the actor has an operative belief that the observer believes to be false, then the observer may say that the actor's belief is a superstition. If the beliefs of the observer are not derived from empirical or inductive procedures, his allegation that a belief is a superstition will not hold up amongst scientists. We, as scientists, are wont to label as superstitions the beliefs that differ from scientifically supported propositions.

How about scientific theories about which there is disagreement? Or how about scientists who persist in believing discredited theories? Should the label ' superstition ' be applied in these cases? In one sense, it certainly should. Beliefs in invalid theories that belong under the institutional aegis of Science are of precisely the same logical and psychological sort as beliefs in the most primitive and mundane ritual.

But this brings up the point of the social psychological significance of calling a belief superstitious. Since the term superstition carries a strong pejorative meaning in our culture, one may run the risk of offending another if he charges him directly with being superstitious. In actual practice it is not the custom to apply the term to individuals with whom some social contact must be maintained, or who are present in person. The term is usually applied to absent persons, to the hypothetical Masses, or to distant cultures. This usage convention involves little risk or retribution, and is properly discreet.

A final point should be made that may seem surprising, in view of the title of this paper. Instead of using the term superstition more frequently, as would be necessary if we were to apply the term at all consistently, it should be used less frequently, and perhaps not at all. According to the major points outlined earlier regarding the psychological nature of superstitions as a category of beliefs, it should be apparent that they are ubiquitous; that even scientists have them though here they are often clothed in the dignified fabric of irreproachable mathematical or verbal symbolism. Moreover, superstitions, i.e. beliefs that are not inductively derived and not empirically checked, are psychologically necessary for existence. Thus, even though the category of beliefs called superstitions is a distinct and meaningful category, it turns out to be too inclusive a class of beliefs to bear the socially reprehensible label of Superstition. For the sake of equity and fairness, the term should be used not more, but less.

Wesleyan University and University of California,
Berkeley

REFERENCES

A. R. Cohen, J. V. Brehm and D. Latane, 'Choices of Strategy and Voluntary Exposure to Information', *J. Personality*, 1959, 27, 63-73.

J. Cohen, *Chance, Skill and Luck*, Penguin Books, 1960.

Elaine Cumming and J. Cumming, 'Mental Health Education in a Canadian Community' in B. D. Paul (ed.) *Health, Culture, and Community*, Russell Sage Foundation, New York, 1955.

F. C. Dockeray and W. L. Valentine. An Analysis of the Elementary Psychology Course at Ohio State University, *J. Appl. Psychol.*, 1935, 19, 503-520.

J. Eaton, 'Folk Obstetrics and Pediatrics : A Case Study of Social Anthropology and Medicine' in Jaco, E. (ed.), *Patients, Physicians,* and *Illness*. Free Press, Glencoe, 1958.

C. J. France, 'The Gambling Impulse', *Amer. J. Psychol.*, 1902, 13, 364-407.

J. G. Frazer, *The Golden Bough : A Study in Magic and Religion* (abridged ed.), Macmillan, London and New York, 1925.

Else Frenkel-Brunswik, 'Intolerance of ambiguity as an emotional and perceptual personality variable'. *J. Personality*, 1949, 18, 108-143.

A. R. Gilliland, 'A Study of the Superstitions of College Students', *J. Abn. & Soc. Psychol.*, 1930, 24, 472-479.

C. S. Hall and G. Lindzey, *Theories of Personality*, Wiley, New York, 1957.

F. Heider, 'Social Perception and Phenomenal Causality', *Psychol. Rev.*, 1944, **51,** 358–374.

W. James, *Principles of Psychology, Vols. I and II,* Henry Holt, New York, 1890.

J. Jastrow, *Fact and Fable in Psychology,* Macmillan, New York, 1900.

R. Lewinsohn, *Prophets and Prediction,* Secker & Warburg, London, 1961.

B. Malinowski, 'Anthropology', *Encyclopedia Britannica,* First Supplementary Volume, London and New York, 1926.

R. K. Merton, *Social Theory and Social Structure.* Glencoe, Ill. : Free Press, 1957.

A. E. Michotte, *The Perception of Causality* (trans. by T. R. and Elaine Miles), Methuen, London, 1963.

J. J. B. Morgan and J. T. Morton, 'The Distortion of Syllogistic Reasoning Produced by Personal Convictions', *J. Soc. Psychol.*, 1944, **20,** 39–59.

W. H. Morse and B. F. Skinner, 'A Second Type of Superstition in the Pigeon', *Amer. J. Psychol.*, 1957, **70,** 308–311.

Jean Piaget, *The Construction of Reality in the Child* (trans. by Margaret Cook), Basic Books, New York, 1955.

Carl R. Rogers and B. F. Skinner, 'Some Issues Concerning the Control of Human Behavior : A Symposium', *Science,* 1956, **123,** 1057–1066.

M. Rokeach, 'Generalized Mental Rigidity as a Factor in Ethnocentrism', *J. Abn. & Soc. Psychol.*, 1948, **43,** 259–278.

J. B. Rotter, *Social Learning and Clinical Psychology,* Prentice-Hall, New York, 1954.

T. R. Sarbin, R. Taft and D. Bailey, *Clinical Inference and Cognitive Theory,* Holt, Rinehart & Winston, New York, 1960.

T. R. Sarbin, Role-Theoretical Analysis of Schizophrenia, Ms. 1962.

L. Saunders, 'Healing Ways in the Spanish Southwest'. In E. Jaco (ed.), *Patients, Physicians, and Illness,* Free Press, Glencoe, 1958.

K. E. Scheibe, 'On the Relationship Between Beliefs and Values, unpublished doctoral dissertation, University of California, Berkeley, 1963.

B. F. Skinner, 'Superstition in the Pigeon', *J. Exp. Psychol.*, 1948, **38,** 168–172.

Smoking and Health, report of the Advisory Committee to the Surgeon General of the Public Health Service, 1964, Washington, D.C.

D. Thistlethwaite, 'Attitude and Structure and Factors in the Distortion of Reasoning', *J. Abn. & Soc. Psychol.*, 1950, **45,** 442–458.

R. H. Thouless, 'The Effect of Prejudice on Reasoning', *Brit. J. Psychol.*, 1959, **50,** 289–293.

E. C. Tolman, *Behavior and Psychological Man,* University of California Press, 1958.

E. B. Tylor, *Primitive Culture* (2 vols.), Holt, New York, 1888.

R. R. Willoughby, 'Magic and Cognate Phenomena : An hypothesis', in C. Murchison (ed.), *Handbook of Social Psychology,* Clark University Press, Worcester, Mass., 1935.

REVIEWS

Method in the Physical Sciences. By G. Schlesinger.
Routlege & Kegan Paul, London, 1963. Pp. vii+140. 21s.

The title of this book promises a wider analysis of the methodology of the physical sciences than the reader finds. In fact, the author limits himself to the discussion of four 'supplementary' principles which he distinguishes from 'primary' principles, such as the law of causality, on the grounds that while the primary principles assert in various ways that there are laws, supplementary principles specify the types of law there are. Primary principles interest all philosophers, but supplementary principles are to be found (it is argued) only by examining the methods 'practical' scientists actually use. The four supplementary principles the author selects for discussion in four essays are those of *simplicity*, of *micro-reduction*, of *connectivity*, and of *verification.*

The essay on the principle of simplicity opens with the author's criticism of Kapp's elevation of Ockham's Razor into the universal principle that 'in physics the minimum assumption always constitutes the true generalization'. Kapp's paper[1] and his subsequent correspondence with Schlesinger[2] were published in this *Journal.* Kapp eventually conceded that in claiming universality for the principle he had claimed too much though 'physicists do in fact often act as though they believed in it'. In rejecting also the universality of Feuer's thesis that 'the verified theory is the simplest because every unnecessary component is an unverified item', the author also displays his own (weakly) verificationist leanings. These leanings become more obtrusive when he next summarily dismisses Popper's arguments for choosing the simpler of two rival hypotheses by means of a single contrived but unconvincing counter-example. Nevertheless, the complexity of the concept of simplicity is revealed by the argumentation.

The author rounds off this essay by introducing a principle of 'dynamic simplicity' according to which a true theory supersedes a rival but false theory because the false theory becomes progressively encumbered with conflicting *ad hoc* elements. He admits that the practical scientist faced with a choice of two such theories cannot hope to be provided with criteria for making a correct choice. Application of the principle of dynamic simplicity requires the scientist statically to wait for one of the theories to emerge—and then dynamically to jump on the bandwagon of emergent truth.

---

[1] R. O. Kapp, 'Ockham's Razor and the Unification of Physical Science', 1958, **8**, 265-280

[2] 1960, **10**, 55-62, and 11, 328-331

The second essay dismisses the principle that the properties of wholes should be explained in terms of the properties of their constitutent simples—the principle of micro-reduction. The author argues that, though the principle has a long history and many achievements in science, the reverse principle—of macro-reduction—has also had some successes. He concludes that the firm belief in the efficacy of micro-reduction, widespread among scientists, will lead towards unification of 'all the sciences' (p. 72), even though the principle itself has no firm logical basis. This seems a bold affirmation to make in the light of a discussion confined to examples in physics and which thus wholly ignores the much more serious reductive problems of the biological and human sciences.

After critical analysis of the relevant views of Mach, Campbell and Bridgeman the author formulates the principle of connectivity, somewhat tentatively, as 'a universal tendency not to allow [physical] systems to differ in what appears to the mind a single disposition' (p. 104). All the author's illustrative examples support the thesis but they are all drawn from classical physics, and the 'open texture' of the empirical concepts of classical physics makes it difficult to delimit 'dispositions' as narrowly as the author's arguments require. Nor is the concept of 'physical system' defined. If, however, atomic electrons are admissible as physical systems, the principle of connectivity at its clearest, strongest and most constructive is exhibited in quantum physics—as the Pauli exclusion principle. But is not the author's tentative formulation of the principle of connectivity merely a weak version of Spearman's neogenetic principles of eduction? And, if so, does it not qualify, on the author's criterion, as a primary rather than as a supplementary principle?

In his final essay the author argues that, though the principle of verification in its strongest form has rightly been discarded as a means of discriminating between meaningful and meaningless empirical propositions, there remains a weaker form of the principle which is useful to the practical scientist. This weaker form is identified with Bridgeman's operationalism—and here the author stops. It seems an odd conflation with which to conclude an essay on the principle of verification. Why retain the word 'verification'? Why retain the principle? And why is falsification not even mentioned in this essay when other and less important matters discussed in *The Logic of Scientific Discovery* (such as the *mathematical* problem concerning the number of points required to determine the parameters of a plane curve of the *n*th degree) are referred to in the earlier essay on the principle of simplicity?

The 'practical' physical scientist to whom this book is addressed will find no discussion of, or any reference to, the exciting methodological problems raised by twentieth-century physics. And I am afraid that he will conclude that the philosophy of science has little relevance to his scientific work.

B. C. Brookes

Models and Analogies in Science. By Mary B. Hesse.
Sheed & Ward, London, 1963. Pp. 150. 15s. 0d.

There can be little doubt that most of the important contemporary discoveries in philosophy of science have resulted from the consequences of a distinction between the context of discovery and the context of justification and a subsequent application of the techniques of symbolic logic to problems which can be formulated in the context of justification. The outcome has been a solution of certain problems of justification which usually depend upon some fully axiomatised theory in which a certain class of theorems (observational sentences) can be objectively assigned truth values as a result of experiment. It has been objected that this approach is not *descriptive* of scientific practice since fully axiomatised theories are rarely used by scientists, and further that this approach is not capable of furnishing any answer to the question of why certain formalised theories rather than others should come to be adopted. These objections are often meant as criticisms of the formalistic approach, but such criticism is misplaced in so far as formalists only claim to be attacking a specific range of problems. On the other hand, an attack on the remaining problems concerned with describing the development of theory in competent scientific contexts represents an important contribution. In the opinion of the reviewer, *Models and Analogies in Science* is welcome as an attempt to use a model notion and an analogy relationship to attack these problems.

The first chapter raises the issues by means of a dialogue between a Campbellian and a Duhemist. This dialogue avoids the trap of making the Duhemist assert that models are *never* useful. The Duhemist agrees that models may be useful in suggesting theories, but he argues that they are not logically essential as an interpretation of a formal system for that system to be acceptable as a scientific theory. Not surprisingly, since this dialogue is of recent invention, the Duhemist is willing to make many concessions because he has quantum theory to play as a trump master. The Campbellian argues that models are essential to the logic of scientific theories if theories are to permit significant *systematic* predictions, and he (or perhaps she) argues against a purely formalist interpretation of quantum theory that classical particle physics is the appropriate observational model. It seems to be true, and the dialogue brings this out, that these positions are not contradictory because *theory* is used ambiguously in their formulation. When a scientist speaks of relativity theory or quantum theory, he seems to be talking about a group of related claims some of which have changed quite radically over the history of the theory. What can make it legitimate to call a changing set of claims by a name which purports to describe a single theory? The Duhemist can reconstruct these changes as a series of adopted and subsequently rejected formal theories, but this raises a rather severe problem of explaining why one new formal theory rather than another should be adopted to replace a rejected one. There are some unresolved difficulties with the notions of falsifiability, simplicity, degree of confirmation, etc., which have been proposed by formalists to solve this problem. From the Duhemists' point of view, these difficulties are obviated by a feeling that any alternative must provide an explanation in terms of myth, or metaphysics. The Campbellite's counterargument is that if a theory is to be *strongly predictive* (that is, able to predict relations between observational predicates not essentially occurring in the data available when it is adopted) its development must be guided by the adoption of an entrenched physical theory as an observational model, and tentative inferences from the model to observational statements of the theory developed by means of an analogy relation. This idea is not new to a Campbellite, but Chapters 2 and 3 of the book explore its consequences rather more thoroughly than it has been explored in past literature. In general, this exploration provides evidence that it may be possible to reason systematically on the basis of a model and an analogy relationship, even if it does fall short of convincing one that this would be the *only* way of systematically guiding the development of a theory. The latter qualification arises from the fact that if one cannot think of a model (there is no guarantee that a model having suitable properties will always be available), one might *systematically* proceed by trying to fit mathematical systems that have had successful application in the past to the available data without regard to a physical model, but this procedure is not compatible with the explicit modelling notion developed in the book.

The definition of model in Chapter 2 depends upon a list of the observational predicates of a model, as well as known causal relations which have been discovered among them. Any predicate of the model may then bear the similarity relationship to a predicate in the observational vocabulary of a theory if it is identical or similar to it. Miss Hesse makes this relationship depend on an intuitive notion of similarity, but it would seem that it could be considerably improved in precision by noting non-causal syntactical properties of the predicates of the model, especially when 2 or more place predicates are considered. (These are not considered in the book.) Thus, if Rxy in the model is symmetrical and transitive (even though neither relationship was essential to the statement of any *causal law* of the model), and these properties would be provable in any suitable axiomatisation of the model, then any Twz in the observational vocabulary of the developing theory would also have to be symmetrical and transitive if the similarity relationship were to hold. An object characterised by the presence of certain observational predicates of a model bears positive analogy to any object having predicates standing in the similarity relationship to these, and the problem of analogical argument is to decide how known positive analogies enable one to systematically predict the occurrence of further similarity relationships between the predicates of the objects having positive anology.

Chapter 3 discusses possible justifications of analogical argument in terms of inductive support, probability, falsifiability, and simplicity. The conclusions are mostly negative, although special forms of analogical argument are thought to be justified by some of the methods. These considerations are not free of reference to implicit axiomatisations of scientific language. For example there is a somewhat puzzling claim (see pp. 93-96) that analogical reasoning can only be successful if a non-Humean notion of causal law is accepted. This is apparently defended (pp. 120-121) by the assertion that although inductive reasoning is possible in a ' chance ' universe, analogical reasoning is possible only if ' stronger-than-chance ' assumptions are made about the universe. But this particular claim is tied to a version of Keynesian inductive inference. Carnap's more careful development entails that neither inductive reasoning nor analogical reasoning is possible in the explicated sense in a ' chance ' universe, while Humean notions of causality *are* compatible with ' stronger-than-chance ' universes in which they are possible. Thus the import of a crucial claim is considerably restricted by its dependence on a particularly controversial technical definition.

Robert Ackermann

The Minds of Robots. By James T. Culbertson.
University of Illinois Press, 1963, pp. 466. $10.00.

' The purpose of this book is to show in detail how states of consciousness can be produced . . . in artificially constructed devices or robots.' Thus boldly does Dr Culbertson stake his claim in the opening sentence of a book which is, whatever else, a remarkable single-handed *tour-de force.* It pursues a line of thought based on the theory of ' artificial nerve-nets ', and expressed in the somewhat daunting language of that esoteric corner of science. It is not the author's first book on the subject, and readers of his *Consciousness and Behavior* (1950) will know roughly the symbolism to expect.

From the outset, Dr Culbertson is clear about the distinction between consciousness, defined as ' mental events or experiences ' and behaviour. ' All present-day robots, computers and other machines are unconscious, no matter how intricate their response may be. . . . No one could tell if a machine were conscious just by observing its behaviour ' (pp. 3, 4). On the other hand, ' mental events consist of certain sets of physical events . . . i.e. neuron impulses. . . . The state of consciousness is a four-dimensional network of impulses such that the constituent impulses do not all exist at the same time, but nevertheless build up into an integrated structure ' (pp. 4, 5).

The gist of his theory is that (p. 212) ' if trees occur, then corresponding subjective phenomena occur ' and conversely. What are trees? ' A tree is very simply described as the set of all relata priorad to any given relatum '

(p. 102). 'Relata' are 'point-events on CTs' and 'CTs' are neuron impulses represented by world-lines in a space-time diagram (pp. 82-83). 'Sense-data consist of . . . CTs connecting together as they pass through a brain, artificial or organic. . . . We will identify consciousness . . . with networks of CTs' (pp. 95, 96).

Does not this mean that consciousness is practically universal? Well, yes 'But the " consciousness " produced in any contemporary computers . . . is so trivially simple that it seems inappropriate to use the term. . . .' If this seems a slightly lame answer, we are reassured in chapter 17 that the theory is testable. Here is one test presented as conceivable in principle (pp. 214-215). An artificial neural net model is connected in one to one mapping to the fibres of the experimenter's optic nerve, in place of his retina. 'If the CT network-theory of perception is true, then the sense data in the psychospaces of the final artificial trees will be experienced by the experimenter, but if he does not observe these sense data, then the theory is false.'

At this point the reviewer must confess bewilderment. The relations between the spatio-temporal figure of excitation of the optic nerve and the visual sensations of the observer could presumably in principle be empirically known before such an experiment. Those between the input to the artificial net and its output to the optic nerve could presumably also be observed or calculated. Assuming that these one-to-one relations were not disrupted by the connection of the network to the nerve, the sensations of the observer must then surely be predictable without reference to any theory of consciousness whatever?

But alas, the reviewer is mistaken. Dr Culbertson eventually makes it plain that on his theory there *is* no one-to-one relation between the spatio-temporal figure of retinal nerve excitation and the visual experience of the observer. What matters is the history (*within the artificial network*) of the events giving rise to this figure. Different input processes producing identical optic nerve impulse configurations could on his theory result in different visual sensations (pp. 216-218 and 379).

Difficulties bristle if one begins to take such a notion seriously. Where, for example, is a line to be drawn between the past history that does and that which does not affect the perceptual equivalence of two identical figures of excitation? Why are we unconscious of so many formally similar 'tree' processes in other parts of our nervous system? and so on. There is, however, a prior question. Does the demonstration that a notion is testable (laudable though this may be) constitute a *sufficient* reason for taking it seriously? In this case, potential readers must be left to form their own judgment.

D. M. MacKay

Anaxagoras and the Birth of Scientific Method. By D. E. Gershenson and D. A. Greenberg.

Blaisdell Publishing Company, New York, 1964. Pp. xv+63. $1.45.

ANAXAGORAS was born in about 500 B.C. in Ionia, went to Athens where he became the mentor of Pericles, and died in exile in Asia minor at the age of seventy. He wrote one book on Natural Science which enjoyed great popularity and was on sale ' at every good bookstall ' in Athens for the low price of one drachma.

Anaxagoras was the first of the Ionian *physiologoi* to come to Athens and to teach there. Like his predecessors and colleagues he was interested in finding out the nature of matter. What are the entities that exist and persist throughout the changes that ceaselessly go on in the world? What are the laws that govern these changes? The authors of this book credit Anaxagoras with having employed two methods of investigation: first, that of observation and, second, that of logical analysis. Anaxagoras is therefore seen here as the begetter of modern scientific method.

In the course of time, all things change and even disappear; but there must always be something that persists, that is, a kind of unchanging, elementary substance. A law of conservation is thus formulated; it expresses the idea that change and destruction are only temporary and superficial. Moreover, the basic units of matter must be small, even infinitesimal, since an infinite variety of objects emerges from any finite thing as time goes on. All objects of finite size contain the same number of units.

The elementary substances must be organic; for life is the fundamental phenomenon. Rather than explaining how life arises from inanimate matter, it is clearly easier to do so the other way round. Inorganic matter is ' denatured ', having lost the complex structure characteristic of living organisms. No perceptible substance is pure, however; this assertion probably means that the basic elements or particles are common to all substances.

There is one force that rules the many, connected phenomena; and this is Mind (*nous*). It is the source of motion. Though real, Mind is not directly perceptible, only through its interaction with matter. But how can something interact with matter when it itself is immaterial? This is a problem that haunts western thought even today. The world is created through the vortex motion that is produced by Mind. At the beginning, things were all mixed up together; but Mind made things revolve and separate out. There is no void, as Parmenides already had said; this obviously fits the world picture of Anaxagoras. The authors end with a brief discussion of some, more specific theories of Anaxagoras concerning astronomy, biology, etc.

This book is the outcome of the co-operation of a physicist and of a classical scholar. The idea of such a collaboration is admirable. For classicists rarely, if ever, understand the scientific content of Greek speculations. A knowledge of what the early, obscure conceptions have become in modern science enables us to see more clearly what the Greeks may have meant. However, regretfully, it must be said that this book does not succeed, all the same. The error of the philologist interpreting the Greek text is to be too literal. The authors fell into the opposite error, namely, of reading too much into Anaxagoras and of trying to make him into a modern scientist. His theory of ' portions ' (*moirai*), though of course a step in the right direction, is not yet on the level of abstraction needed for atomism (and achieved by Leucippus and Democritus). And though Anaxagoras denies the divinity of the sun and the moon, he still has need of a Prime Mover—the source of motion—which he calls ' divine '. (This brought against him the charge of irreligion leading to exile.)

The main criticism I would level against this book is that no single individual can be made into the Hero of Science and that the growth of ideas must be explained as a social-psychological process. The Pre-Socratics, from Thales to Leucippus and Democritus (and not forgetting Pythagoras), must be seen as a group; they lived within a very short period of time and within a very restricted area; and there was quite effective communication and transmission of ideas between them. The few surviving fragments of Anaxagoras' writings seem to stem from the first chapter of his book alone; and we know them only through the quotations of Simplicius who lived in the sixth century A.D., that is, one thousand years later. This makes it impossible, in my view, to interpret the utterances of a single thinker without referring to all his predecessors and, possibly, even to his successors, at least until the time of Simplicius; for the fragments must have suffered distortion. In other words, we can only offer a reasonable interpretation of the whole school of thought, of the movement as a whole. Instead of the intellectualised history in which the words of one thinker are tied to the words of the next, we must try to describe the actual movement of ideas, by uncovering the unconscious processes that underlie all thought. These unconscious processes are brought out into the open by the phantastic misconceptions of the world that the Greeks themselves expressed so openly. Indeed, the very rationality of the Greeks consists in their ability of expressing their phantasies and of criticizing them. Even today, science, that is, any given separate discipline, is still the concern of a relatively small group of creative workers; and there is good, and even quick, communication between them. If we want to understand what makes science ' tick ', we must investigate the group process; and that is the reason why we must always go back to the Greeks, with whom it all started.

E. H. Hutten

ABSTRACTS

Dialectica, 1963, 17, No. 1

A. Pap, ' Neuere Kernfragen der Wahrscheinlichkeitstheorie '

E. E. Harris, ' The End of a Phase '

Recent pronouncements by linguistic philosophers (e.g. Urmson and Warnock) lead to the conclusion that the philosophical foundations of this kind of philosophising have collapsed. It originated in the Logical Atomism of Russell, which, by consecutive steps led first to the rejection of Meinongian objects and then to the repudiation of metaphysics, restricting the subject of philosophy to the analysis of language. Subsequently Wittgenstein's later work unmasked the errors of the still presupposed metaphysical basis. But instead of rejecting the consequent philosophical method, its advocates attribute the defects of the metaphysic, by a viciously circular argument, to the misleading character of linguistic forms. Once this *petitio principii* is recognised, the whole concept of linguistic analysis as the sole method and exclusive range of philosophy must be abandoned.

P. Bernays, ' Remarks to " The End of a Phase " '

E. Emery, ' Propos sur les fondements de la musique '

In order to overcome the present quarrel between the ' ancients ' and the ' moderns ', in which the theorists of music are taking part and in which the evolution of our musical language is implicated, it is legitimate—and indeed necessary—to search for the laws conditioning the practice of music. To this end, the application of an *open method* to the study of the scales affords a particularly fruitful line of approach and is evidence of a process of thought which does not deny science-inspired philosophy; it reveals both the value and the limits of objective knowledge, and enables the specificity of the problem of music to be asserted. Thus musical creation is open to being conditioned by the laws of nature and is subject to social and cultural influences; in the last resort, it depends on the modalities peculiar to the activity of musical consciousness.

M. Maruyama, ' Basic Elements in Misunderstandings I '

ANNOUNCEMENT

FRITZ LONDON'S (1900–54) SCIENTIFIC CORRESPONDENCE, ORIGINAL NOTEBOOKS AND MANUSCRIPTS GIVEN TO DUKE UNIVERSITY LIBRARY

THROUGH the generosity of Mrs Edith London and her family, the Duke University Library, Durham, North Carolina has received the scientific correspondence, original notebooks and drafts of manuscripts of Fritz W. London[1] who, at the time of his death on 30th March 1954, was James B. Duke Professor of Chemical Physics at Duke University. The collection is available in the Manuscripts Department of the Library to scholars and other interested persons; copies of specific items will be furnished on request.

Included in the collection are a total of 1971 different items and 31 volumes. Correspondence, which dates from 1926, with 349 scientists throughout the world includes letter and note exchanges with L. Brillouin, P. Ehrenfest, J. Franck, W. Heitler, W. Keesom, H. Kramers, E. Schrödinger, F. Simon, A. Sommerfeld, E. Teller, G. Uhlenbeck, J. Van Vleck, C. Weyl, E. Wigner, and many others. Early notebooks and draft manuscripts include work on the Heitler-London theory of valence, van der Waals forces, activation energy of chemical reactions, He^4, He^3, a book-length manuscript which was never published, etc. A complete set of reprints of London's published works is available in the Rare Book Room of the Duke University Library.

Use of the collection by scholars and other interested persons is invited. Inquiries should be directed to the Librarian, Duke University Library, Durham, North Carolina, 27706.

---

[1] L. W. Nordheim, 'Fritz London, 1900–1954', *Physics Today*, 1954, 7, pp. 16–17

THE
REVIEW OF METAPHYSICS
A PHILOSOPHICAL QUARTERLY
Edited by RICHARD J. BERNSTEIN

Volume XIX, No. 1 *Issue* 73 *September* 1965

ARTICLES	
Hans Jonas	*Life, Death and the Body in the Theory of Being*
Richard Rorty	*Mind-Body Identity, Privacy and Categories*
Nathan Rotenstreich	*On Shame*
Herbert Hochberg	*Universals, Particulars and Predication*
CRITICAL STUDY	
William L. Reese	*Phenomenology and Metaphysics*
EXPLORATIONS: ARISTOTLE	
Murray Greene	*Aristotle's Circular Movement as a Logos Doctrine*
Norwood Russell Hanson	*Aristotle (and Others) on Motion Through Air*
BOOKS RECEIVED	
Edward A. Reno, Jr. and Staff	*Summaries and Comments*

PHILOSOPHICAL ABSTRACTS
DOCTORAL DISSERTATIONS, 1965
VISITING PROFESSORS FROM ABROAD
ANNOUNCEMENTS

Welch Hall, Yale University, New Haven, Conn., U.S.A.

Readers resident in Great Britain may subscribe at a special rate through our British representative, Professor Stephan Körner, Dept. of Philosophy, The University, Bristol 8, England

---

The Journal of Philosophy

A biweekly journal containing philosophical articles of current interest and encouraging the interchange of ideas. Editors: John H. Randall, Jr., Sidney Morgenbesser, Arthur Danto, and James J. Walsh. $7.50 a year; to students, $6.00; 50 cents a copy.

Now available: Cumulative *Fifty-year Index, 1904-1953:* articles classified by subject and author; 452 p., cloth, $12.00. Also, *Ten-year Supplement, 1954-1963:* $3.00 cloth, $2.00 paperbound. Both *Indexes,* bound together, $15.00.

Back numbers and volumes: 1960 to 1964 (LVII to LXI), 50c/issue, $12/volume; 1940 to 1959 (XLVII to LVI) 75c issue, $18/volume; 1904 to 1939 (I to XLVI), $1/issue, $24/volume.

720 PHILOSOPHY HALL, COLUMBIA UNIVERSITY
NEW YORK, 10027

Philosophy and Phenomenological Research

AN INTERNATIONAL QUARTERLY

Edited by MARVIN FARBER

SEPTEMBER 1965, VOL. XXVI, No. 1

CONTENTS

I. ARTICLES

Intrinsic Value, by MONROE C. BEARDSLEY
Old Mysticism and New Knowledge, by HORACE M. KALLEN
The Mote in the Eye of the Critic of Critical Realism, by W. PRESTON WARREN
Existentialism and Existentialisms, by MIKEL DUFRENNE
The Scale Step, by GERHARD ALBERSHEIM
The Cartesian Circle, by LYNN E. ROSE

II. DISCUSSIONS

On Rose's "Cartesian Circle", by NORMAN KRETZMANN
Brentano on the History of Greek Philosophy, by HUGO BERGMAN
Husserl and Pre-Reflexive Constitution, by RICHARD T. MURPHY, S.J.
The Problem of Pain and Contextual Implication, by THOMAS A. LONG
Lamont on Rights and Duties, by MARCUS G. SINGER
Toward Understanding Aristotle's Categories, by MICHAEL NOVAK

III. REVIEWS, NOTES AND NEWS, AND RECENT PUBLICATIONS

Subscription prices $6.00 per year for individuals; $7.50 per year for institutions. Separate issues, $1.65 and $2.00 respectively. Postage added for back issues only. A circular listing the contents of the Journal since 1940 will be sent upon request.

Address all communications to

PHILOSOPHY AND PHENOMENOLOGICAL RESEARCH
State University of New York at Buffalo, Buffalo, New York, 14214

JOURNAL OF THE HISTORY OF IDEAS

A Quarterly devoted to Intellectual History

VOLUME XXVI — OCT.-DEC., 1965 — NUMBER 4

TABLE OF CONTENTS

ARTICLES

The Eleatics and Aristotle . . . . *A. R. Lacey*
Al-Farabi on Plato and Aristotle . . . . *Majid Fakhry*
Doni's "New World" . . . . *Paul J. Grendler*
Gesture as a Universal Language . . . . *James R. Knowlson*
French Liberalism and Aristocracy . . . . *George A. Kelly*
Dana and Darwinism . . . . *William F. Sanford, Jr.*
Legal Realism and Justice Holmes . . . . *Wilfrid E. Rumble, Jr.*

NOTES

Eve's Dream and Witchcraft Conventions . . . . . . *John M. Steadman*
Political Ideas of Cleghorn, Hume's Rival . . . . . . *Douglas Nobbs*
Edenic Myth in Bancroft's *History of the United States* . . *Merrill Lewis*

REVIEW

Balazs on China . . . . *Conrad Schirokauer*

Subscription Rates (outside U.S.A. post paid): $6.50 for one year; $1.75 (plus postage) for single issue

Payable to

JOURNAL OF THE HISTORY OF IDEAS, INC.
The City College, New York, N.Y. 10031

VOLUME VII JUNE 1965 No. 1

RATIO

edited in co-operation with
H. B. Acton (London), P. Bernays (Zürich), A. Church (Princeton), H. Cherniss (Princeton), G. Henry-Hermann (Bremen), W. C. Kneale (Oxford), E. Nagel (Columbia), K. R. Popper (London) and J. W. N. Watkins (London)
by
Stephan Körner (Bristol)

H. WANG	Russell and his Logic
P. LORENZEN	Methodical Thinking
P. WEINGARTNER	Can one say of definitions that they are true or false?
REVIEWS:	
P. T. GEACH	*Reference and Generality* (L. J. Cohen)
H. N. CASTAÑEDA and G. NAKHNIKIAN	*Morality and the Language of Conduct* (A. Montefiore)
W. H. WALSH	*Metaphysics* (R. Edgley)
E. TOMS	*Being, Negation and Logic* (C. Lejewski)

Price 15s. net **Annual Subscription: 27s. 6d. net**

BASIL BLACKWELL · OXFORD

MIND

A Quarterly Review of Psychology and Philosophy
edited by Professor **GILBERT RYLE**
with the co-operation of Prof. Sir F. C. Bartlett *and* Prof. C. D. Broad

CONTENTS FOR ISSUE OF OCTOBER 1965

Ratiocination not Rationalisation	R. STONE
A Critique of the Ideals of Liberty	H. J. MCCLOSKEY
How Philosophers see Stars	F. B. EBERSOLE
On the Conception of Truth	I. S. NARSKI
"Because I want to"	R. ABELSON
Games, Justice and the General Will	W. G. RUNCIMAN and AMARTYA K. SEN
Are Zenos Paradoxes based on a mistake?	H. N. LEE
Discussions:	
A Note on the Identity Thesis	D. WEISSMAN
The 'Religious Morality' of Mr. Patterson Brown	A. FLEW
Patterson Brown on God and Evil	K. CAMPBELL
Mr. Alston on Hypostatic Analysis	N. GARVER
Mr. Aune on Strawson	R. L. PHILLIPS
Hare on Descriptive Meaning Rules	G. N. A. VESEY
A Reply to Rescher on Goodman's 'About'	T. E. PATTON
A Further Note on Natural Deduction	E. J. LEMMON
Reply to First and Second Thoughts in Moral Philosophy	D. H. J. WARNER and A. C. EWING
Critical Notice	
R. C. Cross and A. D. Woozley, *Plato's Republic*, a philosophical commentary	D. MITCHELL

NEW BOOKS NOTES

Annual Subscription 25s, post free; single copy 7s 6d, plus postage
Published for MIND ASSOCIATION by
THOMAS NELSON & SONS LTD Parkside Works EDINBURGH 9

FOUNDATIONS OF LANGUAGE

International Journal of Language and Philosophy

Editorial Board Morris Halle (MIT), Peter Hartmann (Münster/W), K. Kunjunni Raja (Madras), Benson Mates (Univ. of California), J. F. Staal (Amsterdam), Pieter A. Verburg (Groningen), John W. M. Verhaar (Manila)

VOLUME I No. 2 MAY 1965

Articles J. F. Stall: E. W. Beth 1908-1963—S. R. Levin: *Langue* and *Parole* in American Linguistics — S. Hattori: The Sound and Meaning of Language—R. B. Lees: Turkish Nominalizations and a Prolbem of Ellipsis—W. Motsch: Grammar and Semantics.

Reviews Emmon Bach: An Introduction to Transformational Grammars (Rudolf de Rijk)—Jerrold J. Katz and Paul M. Postal: An Integrated Theory of Linguistic Descriptions (J. F. Staal)
Publications Received for Review

FOUNDATIONS OF LANGUAGE is published quarterly

Annual subscription price *f* 42.—($11,75: 84s.)

R. RIEDEL PUBLISHING COMPANY/DORDRECHT-HOLLAND

Heffers

Current Catalogues from a world famous bookshop

BOOK ADVISOR-NEW/NON-FICTION

HISTORY OF SCIENCE

SCIENCE-NEW/FORTHCOMING

Available free from Catalogue Department

W. HEFFER & SONS LTD

Petty Cury Cambridge

The British Journal for the
Philosophy of Science

VOLUME XVI NOVEMBER, 1965 No. 63

THE LOGIC OF RATIONAL DECISION*

HERBERT A. SIMON†

PHILOSOPHY has long struggled with the problem of deducing one imperative sentence from another. Do the same laws of logic govern such deductions as apply to declarative sentences, or different laws? If different, can the laws of logic applying to imperatives be derived from the laws applying to declaratives?

The Infantry Field Manual of the United States Army at one time contained the following passage:

> Surprise is an essential element of a successful attack. Its effects should be striven for in small as well as in large operations. Infantry effects surprise by concealment of the time and place of the attack, screening of its dispositions, rapidity of maneuver, deception, and the avoidance of stereotyped procedures.[1]

---

* Received 4. i. 65

† The work reported here was supported in part by a grant from the Carnegie Corporation and in part by the Public Health Service Research Grant MH-07722-01 from the National Institute of Mental Health. An earlier version of this paper was delivered at a joint seminar of Western Reserve University and Case Institute of Technology on Rational Values and Decision, 7 October 1964.

[1] *Complete Tactics, Infantry Rifle Battalion*, Washington, 1940, *Infantry Journal*, p. 20. The problem treated in Part 1 is stated, but the solution only hinted at in *Administrative Behavior*, New York, 1947, chapters III and IV, which in turn is based on my doctoral dissertation, *The Theory of Administrative Decision* (University of Chicago, 1942, unpublished). My debt to Jorgen Jörgensen, 'Imperatives and Logic', *Erkenntnis*, 1938, 7, 288-296, will be evident. I owe an equally large, but more recently incurred, debt to Nicholas Rescher, who has kindly allowed me to read an unpublished manuscript, *The Logic of Commands* (London : Routledge & Kegan Paul, forthcoming, 1965) that develops and formalizes the point of view first suggested by Jörgensen. While I start from a rather different end of the problem than does Professor Rescher,

The second sentence, which contains an explicit ' ought ', might be rendered: (1) ' Employ surprise! ' This command appears to follow from the first statement combined with another imperative that is left implicit: (2) ' Attack successfully! ' (3) Success in attacks implies use of surprise.

An adequate logic of imperatives might well permit the derivation of a sentence like (1) from the conjunction of (2) and (3). Similarly, the last sentence of the quoted paragraph, conjoined with (1), could yield such injunctions as: ' Conceal the time and place of attack! ' ' Screen your dispositions! ' ' Manoeuvre rapidly ! ' ' Be deceptive! ' and ' Avoid stereotyped procedures! '

There have been numerous proposals for the construction of systems of logic that would permit such inferences to be drawn. Some of these proposals erect systems analogous to, but independent of, the standard systems of demonstrative logic. Other proposals permit chains of inferences only in the declarative mode, but supplement the demonstrative logic by rules of correspondence between certain declarative and imperative sentences, respectively.

What shall be the test of whether a system of imperative logic is adequate? One possible test would be to compare it with the practices of applied scientists who would seem to have frequent occasions for deriving one imperative from another. What logic does a physician use to conclude: ' Give the patient two aspirin tablets '? How does the engineer arrive at: 'Build the bridge according to Plan A '? And by what reasoning does the management scientist advise the business firm: ' Produce 2,000 units of product next month '?

1 *Normative and Positive Economics*

Implicit in the practice of any applied science there lies some form, or forms, of imperative logic. Let us try to find what this logic is, with particular reference to management science, or applied economics.[1]

---

our conclusions, though distinct are not mutually incompatible. In particular, Rescher's notion of the termination statement of a command corresponds to what is here called the declarative statement obtained by conversion of the command. But while Rescher seeks to construct a separate normative logic, I undertake to show that such a construction is not actually necessary.

[1] Applied economics is especially suitable for this undertaking because the field is characterised by a high degree of formalisation and self-awareness about matters methodological. Thus engineering design is only now beginning to approach the

Normative economics and positive economics (i.e. pure or descriptive economics) live side by side in close symbiosis. Central to the theory of the firm, in its positive version, is the assumption that firms behave so as to maximise profit. But when the applied economist uses the theory as a basis for advising a business firm, he replaces the existential assertion, 'Firms maximise profit', with the imperative, 'Maximise profit!' Often the very same scientist finds himself, on one day, wearing the hat of applied economist, on another day, the hat of pure economist. I have seen no indication that he has any feeling that when he changes his hat he changes his logic—that the rules of inference in the two cases are different. If the profession is faced with a plurality of logics, it is blissfully unaware of it.

The Positive Theory of the Firm

There is no reason, however, why we should limit ourselves to informal tests of this kind. By examining the classical theory of the firm directly we can satisfy ourselves, with as much rigour as we please, as to whether the same or different rules of logic are used when the theory is employed positively and normatively, respectively. The theory is readily put in mathematical form, but for our purposes a non-mathematical version will be adequate.

Stripping the theory to its barest essentials, we assume a firm that manufactures a single product, and sells each month's product during the same period.

1. The total cost (C), incurred by the firm during the month is a function of the quantity of product (X), and of a variable (α), that sums up the effects of supply conditions—wage rates, costs of materials, and so on.

2. The price (P) at which the product is sold is a function of the quantity, (X), and of a variable (β) that sums up the effects of demand conditions—e.g. consumer incomes.

3. We define total revenue for the month as price (P) times quantity (X), and we define profit (π) as the difference between total revenue (PX) and total cost (C).

---

level of sophistication in decision theory that has been present in economics and statistics for a generation. While the examples are carried through in this paper in non-mathematical form, the rigour of the analysis can readily be tested (and, indeed, I have tested it) by making a simple mathematical translation.

4. The positive theory is completed by the assumption that X^*, the quantity actually produced, equals X^M, the quantity that maximises profit.

4′. On the other hand, the normative theory is completed by the injunction: Choose the production quantity X^* that will maximise profit!

To avoid irrelevant detail in our exposition, let us simply assume that there is some unique, finite, positive production quantity for which profit assumes an absolute maximum. Then, the positive economist could derive from (4) and (3) the familiar theorem:

5. Marginal cost (change of cost per unit increase in quantity) equals marginal revenue (change in total revenue per unit increase in quantity) when production is set at the level that maximises profit.

The Normative Theory of the Firm

If we now observe what the normative economist, starting with assumptions (1), (2), (3), and (4′), does in practice, we see that he first replaces (4′) with (4), derives (5), and then obtains from (5) a new imperative:

5′. Set production at the level where marginal revenue equals marginal cost!

In asserting (5′), he appears to have added to the declarative logic a symmetric rule that permits him to replace an imperative with an existential statement, and an existential statement with an imperative. What is the exact nature of this rule?

Note, first of all, that the rule does not permit us to convert *any* equation in the original system into a corresponding imperative. For example, it would not be correct to convert (3) into :

6. Set the production quantity so that profit equals total revenue minus total cost!

Injunction (6) is meaningless, because (3), being a definition, holds identically in the production quantity, hence (6) gives no command at all, unless it be 'Do something!' Statement (4), on the other hand, is not an identity in the production quantity. The remaining three assumptions are satisfied by all nonnegative production quantities; only a subset of these values (in fact, under our assumptions, a unique value) also satisfies (4). It would appear that the conversion of a declarative into an imperative statement is meaningful only if the former is not an identity in the variable to which the command applies in the latter.

Thus far we have had little to say about the environmental variables α and β, that appear in our relations (1) and (2). Suppose we had additional empirical information about one or both of these parameters, say:

7. The supply conditions, α, lie between $\underline{\alpha}$ and $\bar{\alpha}$. Can we convert (7) into an imperative?

7′. Set the supply conditions, α, at some point between α and $\bar{\alpha}$!

We would, of course, regard (7′) as illegitimate, since the variable whose values the command limits is not a variable controlled by the recipient of the command, the firm.

There appear to be two conditions, then, for the conversion of an existential statement into a command. *First, the statement should not be an identity in the command variable; second, the command variable should be a variable controlled by the recipient of the command.*[1] Note that the operator ' ! ', like the universal and existential operators, binds the variable to which it applies—converts it from a free to a bound variable.

The conversion from imperative to declarative mode has not such restrictions. Given any system containing imperatives, by removing the imperative operators from them, we obtain a positive system in which the erstwhile recipients of commands behave as though they were obeying those commands.

A Restatement of the Positive-Normative Relation

Let us now generalise our results. Normative reasoning always takes place in the context of an empirical system. A set of *variables*, each with a range of admissible values, defines the space of logically possible states of the system—the *state space*. A set of *empirical laws*, specifying certain relations among the variables, restricts the *empirically possible* states of the system to a subspace of the state space. The subspace of empirical possibilities has fewer degrees of freedom than the state space. There is a set of *independent variables*, equal in number to the degrees of freedom of the subspace; fixing a value for each of the independent variables is equivalent to selecting a particular, empirically possible, point in the state space.

The independent variables belong to two mutually exclusive and exhaustive subsets: *environmental variables*, and *command variables*. The command variables are those postulated to be controllable by the actor to whose behaviour the commands apply; the others are the environmental variables.[1]

---

[1] Rescher, op. cit. chapter 8, alludes to this consideration but does not provide an explicit rule.

To this system may now be added additional equations, provided these are consistent with the original set. Alternatively, commands may be added on the conditions: (1) that the variables bound by the command operators are command variables, and (2) that when the commands have all been converted to the corresponding empirical relations, the resulting system of relations is consistent.

After conversion of the commands, new equations may be derived by application of the ordinary laws of logic. Any of these new relations that contains command variables may be converted to a command by binding one or more of its command variables with the command operator. From the method of their derivation, it is clear that the commands so derived will be consistent, in any reasonable interpretation of that term, and realisable.

Let us return now to our example of the theory of the firm. In the system described by statements (1)-(3), the state of the world during any time period is specified by the array of values of the variables: profit, production quantity, price, total cost, supply conditions, and demand conditions. For purposes of this example, it is best to consider all six variables as directly observable. The supply variable might be a statistical index of wholesale prices and wage rates; the demand variable, an index of disposable income; the production quantity, a measure of actual production; price, an average price calculated from the actual revenue received; cost, the total amount paid out to employees and suppliers; and profit, the actual change in the amount of money in the till from the beginning to the end of the period. Under this interpretation, each of the three equations states an empirical law—equation (3), for example, denies the operation here of the Scriptural moth and rust.

---

[1] The assumptions stated here amount to defining a causal ordering on the variables that enter into the empirical laws. See Herbert A. Simon 'Causal Ordering and Identifiability', chapter 1 in *Models of Man*, New York, 1957. The use of state descriptions for the treatment of semantic issues in Logic, introduced by Wittgenstein, has been carried furthest by Carnap. For a brief account of his method see Section 18 of his *Logical Foundations of Probability*, Chicago, 1950. Without some such treatment we cannot be sure when or whether particular empirical propositions are mutually independent.

Consider the admissible sets of simultaneous values of the six variables. Statements (1) and (2) each select out a subset (call them A and B) of this set, by restricting the values of C and P, respectively, as functions of the others. These two subsets, A and B, intersect in a smaller subset (call it C) in which both equations are satisfied. Similarly the equation, (3), further restricts the values of the profit, defining a subset, D, of C. Any triplet of values of production quantity, supply conditions, and demand conditions determines a unique element of the subset. Taking the latter three variables as the independent variables of the system, we have the causal ordering:

$$\begin{tikzcd}
\alpha \arrow[dr] & & \\
& C \arrow[dr] & \\
X \arrow[ur] \arrow[dr] & & \pi \\
& P \arrow[ur] & \\
\beta \arrow[ur] & &
\end{tikzcd}$$

Next, we assign each independent variable to an agent, the production quantity to the firm; the supply and demand conditions to its environment or 'nature'. Any additional relation, like (4), that we add to the system, which is consistent with the previous relations, further limits the admissible simultaneous values of X, α and β to a subset of the subset D previously defined. It defines an additional empirical law. If it is not an identity in a variable assigned to a particular agent, and if it can be solved for that variable, then by binding the variable with the command operator we can convert it into a command to that agent regarding that variable. Statement (4′) is such a command.

Although the supply and demand conditions are not mentioned explicitly in Command 5′, it can readily be seen from the derivation of that statement that it expresses the command variable, the production quantity, as a function of those conditions, hence expresses a conditional command. The command becomes absolute when the state of the environment—the actual values of the supply and demand conditions—are known.

Extensions to Uncertainty and Utility

The particular example we have used is extremely simple. No new problems arise, however, for the logic of imperatives in extending the analysis to the much more complex systems that arise in actual applications of normative economics to real-world problems. I shall therefore only sketch these extensions briefly.

Often, the values of one or more of the environmental parameters are not known with certainty to the decision maker. Instead, he may know certain probability distributions of these variables (the case of *risk*), or even this information may be lacking (the case of *uncertainty*). Here, the profit maximising conditions are simply replaced by conditions that the action be selected so as to maximise the expected value of profits, or to minimax profits—to select that action which will maximise profits for the most unfavourable values of the environmental parameters.

In a still further extension, a whole sequence of decisions is to be made, in successive time intervals. The problem cannot be reduced to the previous one, for new information regarding the environmental variables becomes available as time moves on. Hence, the situation describes a dynamic programming problem, essentially, a problem in the calculus of variations.[1]

Another direction of generalisation takes us into considerations of utility, which is a somewhat more complex and subtle criterion than profit. Unlike profits, utilities are not directly observable. Instead, it is postulated that the consumer exhibits a certain consistency in his choices; if we observe him in a sequence of choices—which bundles he prefers to which others, as his alternatives change—we can infer what his utility function is, and what his probability estimates are of outcomes, from his choice behaviour. Again, the rules for conversion to and from imperatives are exactly the same as before.[2]

Normative Side Constraints

Some additional, highly interesting, considerations are introduced in the linear programming models that are so widely used in both positive and normative economics today.[1] Again, let us consider a very simple example: the optimal diet problem. We are given a list of foods, and for each its price, the calory content, and content of each mineral and vitamin relevant to nutrition. Then we are given a set of nutritional requirements, which may include minimum daily intake of minerals, vitamins, and calories, and may also limit maximum intake of some or all of these components.

---

[1] For an example of a normative theory containing the complications of the two previous paragraphs, see Holt, Modigliani, Muth, and Simon, *Planning Production, Inventories, and Work Force* (Englewood Cliffs : Prentice-Hall, 1960). Perusal of chapters 4 and 6, which contain the mathematical derivations, will show that they are precisely parallel to those of the simpler example above.

[2] For a rigorous development of this theory in its modern form, with applications to statistical decision theory, see L. J. Savage, *Foundations of Statistics*, New York, 1954. Savage is concerned, of course, with the normative form of the theory, but there are no evidences that the logic he uses in his derivations is modified in the least by this. For a less technical introduction to normative statistical decision theory, from which the same conclusion can be reached, see Robert Schlaifer, *Probability and Statistics for Business Decisions*, New York, 1959. These are only samples. The literature relevant to this and the previous footnote is enormous, and uniformly consistent with the conclusions reached here about imperative logic.

The diet problem is to find that sublist of foods and their quantities that will meet the nutritional requirements at least cost. The problem can be formalised along the following lines:

1. The total consumption of each nutritional element is the sum of the quantities of that element for each of the foods consumed.

2. The total quantity of each nutritional element must lie between specified limits.

3. The quantity of each food consumed must be non-negative, although it may be zero.

4. Finally, the total cost of the diet (the sum of the products of the quantities of each food by their prices) is to be minimised.

A diet (the solution is not necessarily unique) that satisfies all the relations (2), (3), (4), is called an *optimal* diet. A diet that satisfies the inequalities (2) and (3) (called *constraints*), but which is not necessarily a minimum cost diet, is called a *feasible* diet (and in this context might also be called a *satisfactory* diet.)

Relations (1) to (4) are all declarative in mode. Suppose, for the moment, they have a unique (vector) solution, $\bar{X}^*$. Then, we can make the positive statement that a diet is optimal if and only if the food quantities in it are those specified by $\bar{X}^*$. No unusual kind of logic is involved in arriving at this statement. Suppose we find a (non-unique) vector, $\bar{X}+$, that satisfies relations (1) through (3). Then we can make the positive statement that the diet described by $\overline{X}+$ is feasible or satisfactory.

---

[1] A standard treatise, which works both the positive and normative sides of the street, is Dorfman, Samuelson, and Solow, *Linear Programming and Economic Analysis*, New York, 1958. See especially their chapter 6, which develops the linear programming version of the theory of the firm. Note the casual alternation, which creates no logical difficulties, between the positive and normative modes. On p. 162, just preceding three theorems that follow (by ordinary logic) from the maximisation assumption, we find the authors saying: 'The profit-maximizing firm, then, seeks that feasible program . . . which makes the profitability . . . as great as possible.' On pp. 164-165, they summarise one of these same theorems with the statement: 'If a firm's situation can be described adequately by a linear-programming model, then for the firm to maximize its time rate of profits, it *should* use a number of activities which does not exceed the number of the restrictions that limit its operations.' (Italics mine.)

Now let us view the situation normatively. Equations (1) and inequalities (3) are of the nature of technological constraints, imposed by the environment. The prices are also assumed to be environmental givens. There remain the $N+M$ quantities of foods and nutritional elements, subject to the M linear constraints (1), and the N inequalities (3).

We now add to (1) and (3) two simultaneous commands, (2′) and (4′):

2′. Select a diet X^*, so that the nutritional requirements are satisfied!

4′. Select a diet, X^*, so that the cost is minimised!

Note that the command variables in both cases are the quantities of foods that make up the diet, although the variables that appear explicitly in (2) are the nutritional elements. Hence, we see that, as in the previous example, our designation of the command variables depends on a prior specification of the causal structure of the model consisting of (1) and (3), and a selection of the variables that are assumed controllable by the recipient of the command.

Having postulated the commands (2′) and (4′), we convert them to positive statements (2) and (4), as before, and solve the resulting system for a diet that satisfies all the relations. We now perform the inverse conversion to an imperative:

5′. Select a diet, X^*, that satisfies both (2) and (4)!

The present example differs from the example of the theory of the firm in two respects. First, instead of a single minimisation (or maximisation) command, we now have a whole set of commands, that operate in conjunction: minimise the cost of the diet *subject to* the condition that it meet the nutritional stipulations of (2). Again we see that reasoning about imperatives requires, as context, a complete system of alternatives and consequences; the conversion operation is defined relative to that complete system, which takes into account the causal ordering of the variables as well as interactions among them.

The second new feature in this model is the notion of feasible or satisfactory actions that we obtain by striking off condition (4). Let us now consider the system consisting of the conditions (1) and (3) and the command (2′). In general, this system will have many solutions, rather than an unique one. Suppose that X° is one of these solutions. Is it legitimate to convert to the command (5″)?

5″. Select the diet X°!

The distinction between (5′) and (5″) is, of course, the familiar distinction between necessary and sufficient conditions. If we call the solution, X°, the means, and the relations to be satisfied—in particular, those relations that were converted from imperatives—the ends, then in the case of the command (5′) the means is both necessary and sufficient to the end, while in the case of the command (5″), the means is only sufficient to the end.

Summary of Part 1

Before turning to a more thorough investigation of means-end relations, and of sufficiency and necessity of means, let me sum up this first part of our inquiry into the logic of imperatives and rational decision. No different brand of logic is used in normative economics from that used in positive economics. On the contrary, the logic employed is the standard logic of declarative statements.

Imperatives enter the normative systems through rules of correspondence that permit commands to be converted into existential statements, and existential statements into commands. The conversion process always takes place in the context of a complete model, with the command variables and environmental variables designated, the state of knowledge about the latter specified, and the causal relations among variables determined. In such a context, commands that bind one or more of the command variables can be converted to the corresponding declarative statements; while declarative statements that can be solved for command variables, but are not identities in those variables, can be converted to commands by binding the command variables with the command operator. In the latter conversions, two alternatives are open: the conversion can be allowed always, or it can be allowed only when the decision variables are determined uniquely. The freer rule might be thought to create a danger that contradictory commands will be derived, a question we shall take up in Part 2.

2. *Means and Ends*

The terms 'means' and 'ends' are common in the literature of problem solving, but rarely occur in normative writings on economics or statistical decision theory. Understanding why this is so will prove enlightening. In the models we have been considering, the decision maker is faced with a *given* set of alternatives. In the theory of the firm, for example, he may choose any positive number, X, as the production for the next period. There is associated with each admissible alternative choice a profit, or a utility, or a cost. The decision maker's task is to choose the action that leads to the preferred consequences. This he accomplishes by making a calculation, which may be trivial (as in the simple theory of the firm), tedious (as in a large linear programming problem), or impossible to carry out in practice (as in determining the best move by the minimax criterion in a chess game). The positive theory asserts or implies that the decision maker makes the calculation, or finds the answer without making it. The calculation itself, the process of making it, is not usually a part of the theory.

Consequences and Actions

In the theory of the firm, the quantity produced, X, is, after the fact, an observable comparable to the other five observables of the system. But X, regarded as a command variable, is not an observable at all; it is an action. In fact, we can say that the after-the-fact observed quantity of goods produced is a consequence of the *action* of producing the goods in exactly the same way that the cost is a consequence of that action. It is these consequences that are represented in the state space, not the actions.

Consider the behaviour of a child who is just learning to speak. The child hears the sound 'ahh', say. There is no built-in connection in his nervous system that allows him, on hearing the sound, to perform the appropriate movements with tongue, lips, larynx and lungs that produce that sound. The connection is a purely arbitrary one between certain stimuli, travelling primarily from sensory organs to brain, and certain responses, travelling primarily from brain to muscles.[1]

The sound as heard by the child belongs to what we have been calling the state space; the signals controlling the muscular actions that produce the sound belong to a (different) space of actions. The child's learning consists precisely in using the feedback through the environment to establish a correspondence between actions—and heard sounds. (A deaf child cannot, for this very reason, learn to speak in this way.) Once the correspondence has been established, and uttering the desired sound is no longer problematic, it is convenient simply to

[1] I am well aware of the oversimplification in which I am indulging. However, adding all the second-order corrections to this simplified scheme would only complicate the discussion without changing its conclusions.

ignore the distinction between the action and its immediate consequence, and represent both in the state space.

A situation is problematic whenever there is a weak link in the chain of learned connections between actions and consequences, between means and ends. For the child, the weak link is often very close to the action end of the chain—his problem is how to produce an ' ahh ' sound. Even in the business firm, however, the problem may just as well be ' How do I get the factory to produce at level *X*? ' as ' What is the best level of production? ' The classical theory of the firm assumes the answer to the former question is unproblematic, hence it confounds the action with the resulting production.

The language of classical decision theory, then, describes the world as it is experienced, and refers to actions by identifying them with their most direct and invariable consequences. We must now inquire whether the very simple scheme we used for avoiding an independent logic of imperatives in such a language is also satisfactory when we add a second language of actions, and when we consider the mapping of actions upon consequences to be problematic.

A Language of Actions

Let us first propose a general form for a language of actions, fashioned after modern computer languages.[1] Each statement in the language is an imperative directing that a particular action be performed (e.g. that a machine instruction be executed). A programme is a sequence of such instructions—not, however, interpreted simply as a conjunction of the imperatives, for the sequence defines a particular time order in which the instructions are to be executed, and the specified time ordering may cause instructions to be executed more than once.

The relation between the sequence of commands that constitutes a programme and its consequences is seldom unproblematic. Programming is the art of designing sequences that have desired consequences, and programming almost always requires a more or less extended period of ' debugging ' during which the programme is tested, and deviations between actual and desired outcomes are gradually removed. Though there is a perfectly definite function of programmes on to consequences (expressed in terms of outputs of calculations), the function is immensely complex, and discovering a programme that has specified consequences may be a difficult or impossible intellectual task.

---

[1] Even in computer programming languages, the distinction between an action and its consequences is sometimes systematically confounded. Thus, in algebraic languages like ALGOL and LISP, the name of a function—e.g. ' (X+14) '—may stand for both the programme of actions required to evaluate that function, and for the value itself, which is the output of the programme when it is executed. In programmes as in life, this confounding is limited to aspects of the situation where the relation between action and consequence is not problematic.

Logical Inference in an Action Language

Programme instructions follow one another temporally, they do not stand in a deductive relation to one another. Hence, the rules of programme organisation have nothing to do with the logic of imperatives, as that term is usually understood.

Many programming languages, however, allow for the introduction of *closed sub-routines.* A closed sub-routine is a programme to which a name is assigned, and which can be inserted by use of its name, as though it were a single instruction, in another programme. Thus, in an appropriate language, we might have the programme:

```
Walk one block
    Step with left foot
    Step with right foot
    Halt if end of block, otherwise repeat
```

This programme might be used as a closed sub-routine in another programme:

```
Go to Times Square
    Walk one block
    Halt if Times Square, otherwise—
    Cross street and repeat
```

The relation of 'Walk one block!' to the three instructions that define that imperative, or of 'Go to Times Square!' to its defining imperatives is clearly a relation of end to means.

Now we can raise again the question of inferential relations. From the injunction: 'Go to Times Square!' can we deduce the imperative conclusion 'Walk one block; then halt if Times Square; otherwise, cross street and repeat!'? Patently, the deduction is invalid, for there may be many other ways of reaching Times Square, following other sequences of actions—walking by other routes, taking a taxi, taking a subway, and so on. The programme constitutes a *sufficient* condition for execution of the sub-routine it defines; it is not a *necessary* condition.

Heuristic Programmes

The generalisation that means are sometimes sufficient, but seldom necessary conditions to their ends holds not only when the means are described by a programme (a recipe) but even when the means are described in terms of states of affairs to be realised (a blueprint). Thus, means for bridging Puget Sound were specified by the blueprints for the Tacoma Bridge. Many other designs could have been synthesised towards the same end. As a matter of fact, the Tacoma Bridge did not even prove to be a sufficient means, for under high winds it collapsed. The sufficiency of the means is almost always empirical rather than logical.

Standard design processes employ no new rules of inference from imperatives to imperatives. The design, viewed as a set of imperatives, is simply converted to the corresponding empirical statements. Then the latter, together with the other empirical laws that govern the situation, are used as premises to be tested for consistency with other premises converted in the same way from the design criteria.

Programmes of action that are means to ends only in this pragmatic sense are often called *heuristic programmes.* They are not necessary to their ends, they are not sufficient in any strict sense, they satisfy no particular criteria of optimality. Their claim for consideration is that they frequently accomplish the desired goal, and that when they fail, something else can be tried. They are used because there are relatively few real-life situations where programmes of action having stronger properties than these are available. The map of actions on consequences is seldom simple and unproblematic.

Problem-Solving Programmes

Thus far, we have maintained the assumption that means are given, however problematic their relations to ends. We have talked about how the Tacoma Bridge design could be tested, not how the design might have been arrived at in the first place. In practical life, an even larger time and effort is involved in arriving at possible courses of action than in evaluating them, once they have been designed. Problem-solving activity is activity aimed at both the synthesis and evaluation of performance programmes for accomplishing desired ends.

In problem-solving activity, we have to consider the relations of three sets of imperatives: (A) imperatives prescribing performance goals and constraints; (B) imperatives prescribing performance programmes for attaining the goals (A); and (C) imperatives prescribing problem-solving programmes for discovering and evaluating the performance programmes (B).

In the normative models considered in Part I, no line was drawn between the performance programme (B), and the problem-solving programme (C). In the simple model of the theory of the firm, the optimal action (B) is discovered deductively from the equations of the system, and the method of its derivation guarantees that it satisfies the goal as well as any side constraints that have been imposed. Evaluation is complete as soon as the design has been found.

The same may be said of the linear programming model for the optimal diet problem—with one qualification. In this case, the derivation (C) of the optimal solution (B) is not a trivial matter. If the problem is large, it calls for powerful algorithms and powerful computational means. Under these circumstances, a normative theory of linear programming must include prescriptions of effective computation procedures (C). The theory is incomplete if the calculations are not performable by the kinds of information processing systems—humans or computers—available to perform them. Normative theories are concerned with *procedural means*, processes for finding programmes that will attain the ends, as well as with the substance of the programmes themselves.

We may now ask of problem-solving programmes: is there implicit in them a logic of imperatives different from the conversion scheme that we have already examined? By inspecting the structure of normative and positive theories of heuristic problem solving, we reach the same answer as before. No special logic of imperatives is needed; simply rules for converting imperatives to declarative statements, and declaratives to imperatives.

By way of example let us examine the choice of a strategy in a game of chess when the player is in a favourable position. He can set himself the task of 'designing' or discovering a strategy (performance programme) that, in spite of anything his opponent may do, will lead to checkmate of the opponent. The strategy, if discovered, is a *sufficient* means (B) to the goal of checkmate (A).

As a matter of fact, there exists a systematic algorithm that will discover a checkmating strategy in any position where there is one. It can be shown beyond doubt that strong human players do not use this strategy—and for a simple reason: it would require astronomical numbers of calculations. Nor could any computer be made powerful enough to execute it in a reasonable time.

On the other hand, a heuristic computer programme has in fact been written that is quite effective in finding checkmating strategies.[1] Moreover, this programme parallels, in its gross qualitative features, the programmes used by good human players. Many alternative programmes could probably be devised having the same general characteristics. As in the earlier examples of heuristic performance programmes, the logical relation of means to end in this procedural programme is weak. The checkmate search programme is not necessary to the end of discovering a checkmate strategy. It is not even sufficient; there is no guarantee that if a forced checkmate is actually attainable in the position, the programme in question will discover it. There is only empirical evidence that the programme is often efficient—requires only a moderate amount of search—and no evidence whatsoever that other programmes are not more efficient. The strongest claim we can reasonably make for it is that in a number of situations, and with moderate search effort, it has found checkmates that would do credit to a chess master.

In the present instance the criterion of success is relatively specific—the programme finds a checkmate strategy, or it does not. Thus the performance programme (B) here is logically sufficient for achieving the goal (A), while the procedural programme (C) is only a heuristic means for discovering the performance programme (B). If we examine the structure of problem-solving programmes for other tasks—for example, a more general programme for playing chess—we find that even such specificity is usually lacking. No one has devised a chess-playing programme that looks for 'best' moves, and there is reason to believe that construction of such a programme is infeasible.[2] A programme might, however, play excellent chess by searching for good moves, and terminating its computation as soon as one is found. From all the available evidence, this is precisely what strong human chess players do. Manageable algorithms for discovering solutions to optimisation problems are known for only very restricted classes of problems. If the optimisation goals are replaced by satisficing goals, it often becomes very much easier to discover adequate performance programmes.

---

[1] H. A. Simon and P. A. Simon, 'Trial and Error Search in Solving Difficult Problems: Evidence from the Game of Chess', *Behavioral Science*, 1962, 7, 425-429.

[2] Some programmes choose a good move by determining which move is 'best' in terms of a computable, approximate criterion. But there is no guarantee that such a move is *objectively* best, since the criterion is necessarily very rough. Hence such programmes in fact only choose good moves—if they play well at all.

The problem-solving process is not a process of ' deducing ' one set of imperatives (the performance programme) from another set (the goals). Instead, it is a process of selective trial and error, using heuristic rules derived from previous experience, that is sometimes successful in *discovering* means that are more or less efficacious in attaining some end. It is legitimate to regard the imperatives embodying the means as ' derived ' in some sense from the imperatives embodying the ends; but the process of derivation is not a deductive process, it is one of discovery. If we want a name for it, we can appropriately use the name coined by Peirce and revived recently by Norwood Hanson: it is a *retroductive* process. The nature of this process—which has been sketched roughly here—is the main subject of the theory of problem solving in both its positive and normative versions.

Conclusion

This paper has undertaken to throw some light on the logic of imperatives by examining areas of scientific activity where imperatives enter into the scheme of things. Part 1 examined the classical theories of rational decision that have developed in economics and statistical decision theory. Part 2 examined the logical relations of means and end in various programming and problem-solving structures.

In all cases we have seen that the logic used in positive theories is exactly the same as the logic used in normative theories. Speaking strictly, there is never any occasion to deduce one imperative from another. There is frequent occasion, however, to convert an imperative to a declarative statement, and adjoin it to a set of empirical statements. There is occasion, also, to convert a derived empirical statement to an imperative, by binding a command variable. The conditions for carrying out these conversions have been stated.

The final paragraphs of Part 2 show that there is also frequent occasion to obtain one set of imperatives from another set by processes of discovery or retroduction, but that the relation between the initial set and the derived set is not a relation of logical implication.

Carnegie Institute of Technology,
Pittsburgh,
Pennsylvania,
U.S.A.

CAUSALITY, CAUSAL LAWS AND SCIENTIFIC THEORY IN THE PHILOSOPHY OF KANT*

GERD BUCHDAHL

I

It is well known that in the Second Analogy Kant purports to give a proof (transcendental proof) of the causal principle, which he variously denominates 'Principle of Production' (1st edn.) and 'Principle of Succession in Time, in accordance with the Law of Causality' (2nd edn.); the actual statement of the principle in the first edition being:

> Everything that happens, that is, begins to be, presupposes something upon which it follows according to a rule.

The second edition statement runs:

> All alterations take place in conformity with the law of connection of cause and effect. (A189=B232)[1]

It may be worth while to recall Kant's definition of a causal relation, or 'the concept of cause'. (Reference to the chapter on 'Schematism' (A144=B183) shows that this is actually the definition of the schematised concept, the category being the concept of dependence, or ground and consequence.) At A90=B122, this runs:

> The concept of cause . . . signifies a special kind of synthesis, whereby upon something, A, there is posited something quite different, B, according to a rule.

Or A91=B123:

> This concept makes strict demand that something, A, should be such that something else, B, follows from it necessarily and in accordance with an absolutely universal rule.

And we may also compare this with A112-3:

> The concept of a cause involves the character of necessity, which no experience can yield.

It is worth while to remember these definitions for they will prevent our reading 'uncritical' overtones into Kant's conception of

*Received 27. iv. 65

---

[1] *Critique of Pupe Reason*, trans. Norman Kemp Smith. London, 1933.

causality in general. Causality for Kant has to do solely with determinate succession in time 'in accordance with a rule'.

Here one may ask: 'In accordance with what rule?' And the answer which suggests itself might perhaps be: some rule or other, to be specified, to be obtained from experience. This is indeed a possible answer, but I shall want to argue that it cannot be made to square (when understood in an obvious way) with at least one interpretation of the purport of Kant's general argument in the Analogy, although there are passages which lend a certain plausibility to it.

It must be admitted that there is something in the feeling that the kind of answer I mentioned is the correct one. And we can easily sketch a context in which it *is* perfectly permissible. For this purpose I will take the case of a law of nature which is not actually causal, but involves functional dependence and which Kant himself seems to mention at least by implication, in the *Prolegomena*, par. 20, viz. Boyle's Law, although its *application* can be worded causally. Thus, if I apply an increased pressure to a gas, say by pushing down the piston in a cylinder, this will be accompanied (without an intervening time lapse) by a diminution in volume. Now that this relation is 'causal' in Kant's sense, can be expressed by saying that the pressure and volume changes take place in accordance with a universal rule, the rule that the product of pressure and volume at each state of the process of compression (provided the temperature remains constant) is constant, the rule being normally expressed mathematically as pv=const.[1]

Now what is correct, is that the notion of 'rule' which occurs in Kant's definition of causality has as its *model* these kinds of cases. Indeed, for Kant, 'conceptual thinking', and thinking in accordance with rule, are often identified.[2] So when Kant is going to argue that for anything to become an object for a judgment of experience it must involve subjection to rule, this—whilst employing the concrete cases as models—must be understood as being a perfectly general statement, implying no more than the need for conceptual thinking. If so, there is no logical need to give any reply to the question: 'what rule'?, when discussing the general principle of causality. That principle can be regarded as standing so to speak on its own feet; to make clear its bearing in the *Critique*, it is not necessary to complete its meaning by adding: 'And the particular rule involved needs of course to be specified, e.g. through empirical determinations.'

---

[1] It is worth noting in parentheses, that although this is an instance of a deterministic rule, this need not be so even though Kant may not have envisaged any alternatives. If one admits the concept of statistical laws, a statistical rule would be quite feasible, and perfectly consistent with this definition of causality.

[2] cf. at A126 understanding is characterised as 'a faculty of concepts', as well as 'faculty of rules'. Again, a line is viewed (whether as concept or as object—the two are of course eventually made to coincide in the transcendental deduction) as something 'drawn in space', which is regarded as 'a determinate combination of the given manifold', itself constituting a kind of 'unity' (B138), a unity which is in turn said to 'act as a rule' (A127).

The point of that denial brings me however straightaway to the core of the matter, which is, What is the relation between the Kantian Principle of Causality and the special causal laws, the laws of nature? Now it is my contention that Kant does not have a single, univocal view, but that he embraces at different stages of his writings a variety of doctrines, not always mutually consistent. On the other hand, it is of importance to see the relationships between these different views, for an awareness of the variety of slides in which we find Kant's argumentation involved can help to throw into fresher focus each of the separate contentions; and illuminate in general the place of causality both in Kant, and in those who differ from him.[1]

2

As a start, it may be as well to indicate the different levels to which I have alluded, corresponding so to speak to different strengths of causal grip on the phenomena. I will first list them in the order of increasing strength, i.e. starting with the weakest. It should be understood that the 'weakest' means here also the most general, and the most pervasive version.

1. The argument of the Second Analogy, e.g. the part involving the example of the ship moving down the stream (A192=B237), and the corresponding position in the Transcendental Deduction (2nd edn., B142).

According to this argument, to which I shall return in more detail below, the principle of causality is a condition of objective experience in general. The necessity posited as holding between my 'representations', holds for 'empirical' and hence 'contingent' judgments in general. It can hence not be regarded as relating causally the separate parts of an 'event', in the sense of a sequence of states, neither in a particular, nor in any general causal manner.

---

[1] It is of course perfectly possible that these different positions are due to a loose or over-technical expression of the logic of the situation, and that Kant does really wish to maintain only a single view-point. If this is the case, I should want to argue that the view-point which best squares with the rest of his philosophy is the one with which I shall begin my list.

It follows that to say that the Principle of Causality is a transcendental condition of experience is not the same thing as to say that nature is lawlike in general, if this locution is understood in its usual sense, e.g. in some such sense as occurs in Mill. In other words: the Kantian causal principle is not a principle which can be employed as a major premiss, and whose instances are causal laws.

2. A slightly 'stronger' version occurs at B165, where Kant remarks that the understanding is not in a position to prescribe to nature more than causality in general, 'that is, . . . the conformity to law of all appearances in space and time'. And he goes on:

> Special laws, as concerning those appearances which are empirically determined, cannot in their specific character be derived from the categories, although they are one and all subject to them. To obtain any knowledge whatsoever of these special laws, we must resort to experience.[1]

Here, it seems suggested that Kant believes himself to have shown that nature is lawlike in general; i.e. that the notion of empirical judgment of some matter of fact involves the conception of that fact being lawlike in general, and not just being brute contingency. I shall consider this argument together with the next.

3. There is a third version, I am not sure whether it is stronger or weaker than the second; it occurs in connection with Kant's attempt to settle accounts with Hume, in the first chapter of the 'Transcendental

---

[1] It is interesting to note how frequently even now-a-days a number of writers forget this proviso, when commenting on Kant's doctrine of science. Part of the trouble is perhaps that Kant refers to the principles of substance, causality, etc. as 'laws of nature', e.g. at *Prol.*, par. 15; where however it is made perfectly clear that the reference is to these general principles alone. An additional source of confusion is the celebrated passage in the Preface to the second edition, Bxiii–xiv, where it seems to be implied that the scientist must approach nature with his 'principles in hand', the suggestion being that these principles are ordinary empirical laws. Now this would indeed accord with much recent writing in the philosophy of science, both of those who have emphasised that scientific logic does not amount to a generalisation from experience, but the testing of hypotheses, entertained prior to the investigation; and even more, with the views of those who entertain a 'conventionalist doctrine' in some form or other, at least concerning the fundamental principles of a science like dynamics or thermodynamics, or quantum mechanics. Moreover, as we shall eventually see, the latter standpoint accords very well with Kant's treatment of the relation between his transcendental principles and certain scientific laws, such as the laws of conservation of mass, of inertia, and of momentum.

Doctrine of Method '. The gist of it is as follows. Hume was correct to argue, according to Kant, that one cannot *a priori* determine some particular cause (such as that of the action of the sun's rays) being responsible for some given effect, e.g. that of melting a piece of wax. But Hume inferred wrongly from the contingency of the determination of this particular sequence *in accordance with the law* of causality, the contingency of the law itself (cf. A766=B794).

Now this can be read in two ways. One interpretation would give us the ' strong ' version (2); the other would be consistent with the weak version (1). We get the strong account if we imagine that Kant is saying that the law of causality is just the general connective which must be presupposed as supporting the special law here referred to or, rather, the special causal relation, the sun's heat causing the melting of wax. This is perhaps supported by the phrase: ' in accordance with the law of causality.'

However, a second interpretation is possible. Kant may be saying two things. (1) He has proved the law of causality, in the sense that he has shown that the concept of a causal connection in general (' determination of representations in accordance with a rule ') is transcendentally involved in the notion of empirical judgment. So he has given an *a priori* proof of the law. (2) That principle is now just the principle, in accordance with which that special sequence is *interpreted,* is formulated. The difference from the first case will be clear. There we assume that it had been shown that events are lawlike, and it only remains to determine which causes what. Whereas in the second case, it is only the *conceptual employment* of the causal principle which is injected into the situation.

This interpretation is moreover made very graphic by the use Kant makes of just this example in the *Prolegomena,* though once again using a language which suggests the strong rather than the weak interpretation. Yet what is in fact there shown is only that without the causal *concept* certain transitive or accusative states of affairs could not be rendered syntactically in our language. Thus, at *Prol.,* par. 20, first footnote, we read:

> When the sun shines on the stone, it grows warm. This judgment, however often I and others may have perceived it, is a mere judgment of perception, and contains no necessity ; perceptions are only usually conjoined in this manner. But if I say, ' The sun warms the stone ', I add to the perception a concept of the understanding, viz. that of cause, which connects with the concept of sunshine that of heat as a necessary

> consequence, and the synthetical judgment becomes of necessity universally valid, viz. objective, and is converted from a perception into experience.[1]

Here it almost looks at first sight as though Kant was maintaining that the addition of the category concept unaidedly converts the uniform association into a necessary and specific connection. That would be an even stronger claim than any so far made, but I think that it is so flagrantly inconsistent with everything else we have heard, that it cannot be taken *au pied de la lettre*. There are various ways of coping with it. One is this: Kant is explaining by means of an empirical analogy or model, how the category (when applied to the sequence of 'perceptions') converts a 'merely contingent', or 'acccidental' sequence,[2] into a necessary one. And just as this happens, conceptually, at the empirical level, so it does at the transcendental level of the argument of the Second Analogy (as will be explained presently). In other words, Kant is explaining at the empirical level the relation between pre- and post-categorised 'experience'.[3]

But into this plays also a second strain. For it is of course quite true, that without injecting the concept of causality, we could not assert a proposition like 'The sun warms the stone'; this is a causal proposition and not any number of uniform associations can properly yield a causal proposition.[4] But the objection might then be pushed: It is one thing to show that the analysis of causal statements involves a concept additional to that contained in statements of constant conjunction. It is quite another thing to claim that the necessary addition of this concept gives us any grounds for believing that the causal proposition has thereby been justified in the way Hume demanded. That the causal concept involves necessary connection, Hume was the first to admit. That we have this concept, he also agreed. That the concept is involved in all statements of the causal kind, must again be admitted. But unless it can be shown that the causal *principle*, which is instantiated in special causal relations, has a justification, all this does not avail us anything.

---

[1] The distinction between judgments of perception and of experience does not of course verbally appear in the *Critique*, and has not been accepted by many commentators, e.g. Cassirer, Kemp Smith. But we can ignore this here, and simply concern ourselves with the contention that without the causal concept it would not be possible to assert that the sun warms the stone.

[2] The significance of these terms will be made clearer in the next section.

[3] This can also be seen from a parallel example that occurs in the Transcendental Deduction, B142, where the subjection of a series of representations, said to be 'only subjectively valid', to the objective principles which embody the categories, converts that series into an 'objectively valid judgment'. But the example mentioned is 'This body is heavy', which is expressly said to be an 'empirical, and therefore contingent' judgment.

[4] This has after all been made abundantly clear by recent discussions concerning the logical analysis of lawlike statements. For it has been shown that lawlike statements do seem to involve more than mere universals of fact. And this would be just the same point, couched in more formal language.

Now as we saw, Kant did of course claim to have given a justification. But the question is: how does his justification bear on the present case?

I think, it is quite possible that Kant would have refused to take these objections seriously, simply by saying that he was here being asked to give a kind of justification that according to the principles of his philosophy it is not possible to give. This would indeed be the import of his reiterated caution, that the principle of causality had only been proved as applying to mere appearances, to phenomena, and not to things-in-themselves. People have a habit of quoting this proviso, without, it seems to me, quite grasping its import. For it might quite well amount to the contention that the principle did only involve the weak claim (1), mentioned above. In which case Kant could not be reproached with having thought that he was giving a Hume-Mill-like support to the special laws of science. All he would then be taken to claim is the following: The law of causality is *a priori*. The concept of causality cannot be generalised from mere uniformities. Special laws are *formulated in accordance with* the principle of causality, as stated. Which is not at all the same as to say that the principle of causality *supports* these special laws, in the way in which a Mill would claim that his special causal laws are *instances* of his general law of causation, itself a generalisation of these instances. Nor would the relation be like that which exists between such presuppositional principles as those of Limited Variety, Atomicity, and Spatio-temporal Indifference in the inductive philosophy of J. M. Keynes. (The Kantian case would be better represented by the example of the relation between the addition of quantities of a physical kind, in accordance with the rules of arithmetic.)

It will be clear that my different interpretations hinge very much on the argument that the Second Analogy involves the 'weak' doctrine of causation. We must therefore look at the Second Analogy in slightly more detail.

3

The crux of this argument involves an analysis of what we are to understand by, and of how we can frame, empirical judgments concerning a given objective succession of the states of a thing. This analysis is of course carried out in the light of the general philosophical position lying behind the argument of the transcendental deduction.

Suppose we are given two such states, as for instance (to take Kant's own example) the positions of a ship moving downstream, A, followed by B. What makes the cognition of such a change possible? To ask, what makes it possible, is for Kant to ask, what elements are presupposed in this situation, and in which logical realm are these elements located?

The first element is that of sensory perceptions, regarded as mere ' content, or matter of sensation '. Now the central contention of the *Critique* is that this sensory content (in its *a posteriori* aspect) must be viewed so as not to ' contain ' the objectivity of its time succession. Indeed, even the fact of its existing in time *in general* is to be regarded as an *a priori* function, as an aspect of ' mind ', viz. the *manner* in which the *a posteriori* content comes to affect the mind. This ' manner ' relates to the way in which perceptions are ' connected, ordered and related ', or rather: the *possibility* of such a relating. In so far as the relatedness has reference to ' time ', Kant expresses this by saying that the empirical manifold is ' always generated in the mind successively ' (cf. A99; A190=B235; A198=B243). In this way, perceptions are set in a necesssary relation to, and dependence upon, the ' possibility of experience ', and the process of knowledge and consciousness. They are, so far, no more than ' inner determinations of our mind in this or that relation of time '; the model used is that of the subject itself placing the perceptions in time (A197=B242).[1] (In order to indicate this aspect of the ' subjectivity ' of my perceptions, prior to their ordering in accordance with categorical concepts, I shall write them thus A′, B′.)

However, the analysis of the content of a judgment like ' A is followed by B ' is not yet complete. It involves (according to Kant) the further feature of A′, B′ being *combined* in a single consciousness. Once again, the model through which to express the 'subjectivity' of this, is to say (using Kant's actual words), that 'apprehension is only a placing together of the manifold of empirical intuition'; and that in this synthesis, so far 'perceptions come together only in accidental order' (B219). Now this clearly still does not exhaust the full meaning of the judgment: 'A is followed by B.' For in that judgment I do not merely say that I have placed A′, B′ together in an 'accidental' order with respect to the general time relation. Rather, in addition, the judgment 'B follows A' contains the *thought* that the temporal positions of A′ and B′ are not as I, the 'observer', have 'constructed them in time', but as they 'exist objectively in time' (B219).[1] The notion of A being followed objectively by B thus involves the additional thought of objective time determination. Since to demand that the notion of such an objective time in general (e.g. the objectivity of a succession) should represent something that can be 'perceived' is nonsensical—or rather: not permitted by the argument against things-in-themselves—it must be replaced by some logical simulacrum, which relates these perceptions in time in general; according to Kant some general concept of time-determination. In the case of succession, this means that we have to *think* the 'subjective' or 'accidental' succession A′—B′ as being tied down in a determinate way; this is done by thinking it in accordance with a concept which will assure that A is *essentially* placed before B, and never *vice versa*. Such a concept is the concept of a relation such that when we are given a perception, say B′, it is preceded by another perception, here A′, in accordance with a universal rule. But this concept is, with Kant's definition, the concept of causation.

---

[1] The reason for this move is fundamental, being the principle that the aspect of ' objectivity ' of the objects of experience is not to be regarded as that of ' things-in-themselves '. This requires that the concept of object shall become a subjectively spontaneous, transcendental function. It follows that it can be added only to an order which is purely ' subjective '. (Naturally, we are only speaking of an order in general; *what* is ordered, is an *a posteriori* matter.)

To sum up: The subjective succession A′—B′ becomes an objective succession A—B, by adding to the former a conceptual determination; by regarding A′, B′ determined by the causal relation.

Now for the purpose of this paper, the vital point is this: A—B is an objective succession. If my characteristation of Kant's argument is correct, there is no suggestion that A—B is a causal succession. On the contrary, it must not be so regarded, otherwise the proof would be weakened. Remember that the category of causality is to be proved as a condition of experience in general. Now there can be no doubt that we have such experience; but it would be extremely tenuous to require that experience itself to be causal!

---

[1] In the 'Transcendental Deduction' Kant argues that 'experience' and 'empirical consciousness' are tantamount to the formulation of such judgments, and hence only possible through proving the possibility of the latter (cf. A103-104 ; B138).

Rather, so that A, B can be an objective succession, it is A′, B′ that has to be regarded as related under the concept of a necessary connection, or rather, a causal connection, which means succession in accordance with an absolutely universal rule. But what is the status of A′, or B′? Well, it is what Kant calls a 'perception' (Wahrnehmung), which is an element of the real world regarded as a modification of the mind, as purely subjective. That same element, A′, in order to be a member of an objective cognition, to become an A, must be hooked up to a B′, to yield the possibility of such a cognition as A—B.

There can be no doubt therefore that A—B is an objective, contingent, succession, the statement I use to refer to it, a contingent statement. Nor will a variety of manoeuvres usually attempted at this point, be of any avail,—not to mention Schopenhauer's reading according to which Kant has assumed that the ship *must necessarily* go downstream, as a matter of law. One attempted reading is as follows:

In regarding A, B as A′, B′, I regard them as subject to determination in accordance with some rule or other. And it is up to experience to help me find out what that rule is. For instance, in the crudest case, the succession A—B might well be an instance of a generalisation. But often the relation may be more complex, and the causal connection between A and B occur *via* a number of intervening factors.

But this would not be to the point at all. For the transcendental proof *depends* on regarding A—B as an absolutely contingent empirical, indeed, singular statement. The ingredient of causality which we have shown to pervade the situation can hence not be used up as support for some putative causal connection of which our succession is an instance or a relative part.

At this point we may perhaps return to the observation which I noted at the start: the problem of what answer to give to the question: 'What rule?' For an objector may perhaps urge the following: To cognise the objective succession A—B, I have to think A, B (which, quâ appearance, is regarded as A′, B′) as being subject to rule. And this entitles me to try and discover the rule involved, albeit by experience.

There are two comments to be made on this move. (1) It is of course perfectly true that Kant's proof involves the contention that A′, B′ are subject to rule. But the actual laws that could be discovered, concerning the embedment of A′, or B′ in a network of law and theory would be laws which have not the slightest relevance to the original relation A followed by B, except in so far as they regard the latter as an objective contingent sequence. But it was in order to cognise *this* sequence that I had to think A′, B′ as subject to rule. The A′ must be regarded as necessarily connected with the B′, subject to rule. But we are not to assume that the concept involved says: A′, B′ must be subject to some law or theory or other, somewhere in the realm of nature.

2. If this reading is right, then there would seem to remain one perfectly obvious procedure. We could say: We have shown that the statement of the principle of causality can be proved *a priori*, in the sense that the conception of an objective world entails the thought of the members of that world, *regarded as appearances*,[1] being subject to rule. Let us then use this principle, not considered as a *support* for the specifications of the particular rules which experience may come to discover, but as a guide-line for the search of principles; where having this guide-line does not mean having a guarantee that any putative principle *is* a law, but rather, that it gives us the *language* in terms of which to carry on our procedure.

4

Now Kant has a perfectly good terminology, through which to characterise this employment of the categories when applied to the realm of phenomena, e.g. objective sequences of the kind we have in our example of the ship, A—B. The term is ' regulative principle of reason '. What is the relation of the principle of causality to the phenomena, when regarded as a regulative principle of reason?

If we are correct in our interpretation it ought only to *direct* us to search for causes, to formulate causal laws; moreover if we want to have a natural science, we can do no other. Now this is precisely what we find. At a crucial place in the argument of the Antinomies, where Kant seeks to reconcile universal natural causality with the possibly 'fictional' (A545=B573) assumption of moral freedom, he indicates the function of the principle of causality as follows:

---

[1] ' regarded as appearances ' : this locution of course contains the fatal ambiguity, the slide, from A′, B′ to A, B. To regard an object as ' appearance ' is to regard it as being ' subjective ' so as to be in need of the addition of the objective time relation. The ' appearance ' contains the ' ingredient ' A′, B′ ; the addition of the time relation yields A, B. It is therefore systematically ambiguous to conclude the argument by saying that ' the appearance is subject to concepts which supply a necessary connection '. The Kantian terminology which conceals this slide has ' apprehensions ' and ' perceptions ' for what I have labelled A′, B′ ; and ' appearance ' for A, B ; with the plural ' appearances ' (also ' manifold of appearance ') straddling either ; and ' representation ' being a neutral term which likewise covers each (cf. B219, B234 f.).

> The principle of the causal connection of appearances is *required* in order that we may be able *to search for and to formulate* (suchen andangeben) the natural conditions of natural events, that is to say, their causes in the field of appearance. If this principle be admitted, and be not weakened through any exception, the *requirements* of the understanding, which in its empirical employment sees in all happenings nothing but nature, . . . are completely satisfied; and physical explanations may proceed on their own lines without interference. (My italics.)

Here the contrast is made quite clear. The understanding sees in its empirical employment nothing but nature: this is Kant's way of saying that in order to cognise any empirical contingent state of affairs, I have to think the causal concept injected on each and every occasion; which is tantamount to the proof of the causal principle, in this sense of 'proof'. On the other hand, when we turn to a 'scientific' study of nature, this principle is *required* in order to determine causal laws; it is *needed* so that we may search after such laws. Now of course, this passage is itself ambiguous; it may mean: without the causal principle, no law would be *justified* quâ law. Or it may mean: without the causal principle we would not have a *language in terms of which to search for* laws; both interpretations we have noted before. In other words: in order to do science, you have to pretend that your task is to fit causes to natural effects. This latter interpretation is suggested by what Kant himself says about the function of regulative principles of reason; to wit, that the search for conditions can only be 'set as a task', since the total of conditions is not given (A508=B536); that such a principle is purely 'doctrinal' (A516=B544), which for Kant means: methodological. We know of course why Kant had to hold to such an interpretation. Without it, as he points out, we should for instance involve ourselves in the contradictory theses sketched in the Antinomies. In other words, the regress from conditioned to unconditioned must be constrained; and Kant constrains it by making the application of principles like that of causality a regulative matter.

If this reading is right, a number of things Kant says, even in this section, cannot exactly mean what they seem to say and need a certain amount of reinterpretation. For instance, at A542=B570, we read:

> That everything which happens has a cause, is a law of nature. . . . Consequently, all events are empirically determined in an order of nature. Only in virtue of this law can appearances constitute a *nature* and

> become objects of experience. This law is a law of the understanding, from which no departure can be permitted, and from which no appearance may be exempted. To allow such exemption would be to set an appearance outside all possible experience. . . .

But what we have to see is that this kind of language does not help us in our search for an interpretation. In virtue of the law of causation appearances constitute a *nature*. Now by ' nature ' Kant means ' the aggregate of all the objects of experience '; moreover, in such a way, that the existence of these objects is thought of ' in conformity to law ' —e.g. the principle of causality (*Prol.*, pars. 16, 14). So even here it is only claimed that in virtue of the law of causality, and the other principles, do we have a realm of *objects*.[1] On the other hand, the phrase ' determined in an order of nature ' is normally taken by Kant to denote a deterministic system of theory, as we shall see presently. It follows that when Kant here contrasts freedom with determinism this cannot mean what that expression normally suggests; it refers us merely to the contrast between appearance regarded as necessarily subject to rule, and something else, not regarded as appearance, and hence not subject to the concept of rule. Whereas the empirical contrast *normally* intended by this phrase is that between a universe subject to a system of empirical lawlike determinism, as against one in which law does not act. What Kant has done is to sharpen the contrast, by opposing two much queerer conceptions, this being the contrast concerning which alone anything *a priori* can be asserted.

5

A glance at the Kantian conception of ' empirical law ' will help to clarify the whole situation considerably. We have already seen that Kant never doubts that only ' experience ' is available for the discovery of such laws, where ' experience ' includes a reference to the complex of the inductive logic of science as well as the ' method of hypothesis '.[2]

---

[1] It is true that the phrase ' object ' contains itself a similar ambiguity, at the level of the argument of the Analytic. It may have reference simply to the possibility of ' experience ', where this just means ' cognition ', including for instance dreaming, imagining, etc. Or it may suggest something rather stronger, e.g. what is phrased in ' reproducible scientific observation reports '. A clarification of this ambiguity would have to proceed on lines similar to those suggested in this paper, and there is no need to pay to it any further special attention.

[2] Kant's references to scientific methodology in the *Logic*, as well as the first *Critique*, e.g. A789-792, show that he had a perfectly clear (if rudimentary) picture of inductive and hypothetico-deductive logic.

Now these laws, being only 'empirically known', are 'as regards the understanding, contingent'.[1] And not only the individual laws, but that whole 'system of laws', with which we operate when incorporating them into a theory, is contingent likewise.

> We must therefore think in nature, in respect of its merely empirical laws, a possibility of infinitely various empirical laws, which are, as far as our insight goes, contingent (cannot be cognised *a priori*), and in respect of which we judge nature, according to empirical laws and the possibility of the unity of experience (as a system of laws), to be contingent.[2]

However, despite this contingency, laws must be conceived, quâ laws, 'as necessary, for otherwise they would not constitute an order of nature; although their necessity can never be cognised or comprehended' by the understanding.[3]

Here again it is important that we should clearly distinguish the contingency and (!) necessity of empirical laws from the 'synthetic *a priori*' status of the causal principle. There is only a broad analogy between the two, an analogy corresponding to the levels of understanding and reason in the Kantian architectonic. For just as the necessity of the causal principle is generated by the requirements of the transcendental unity of consciousness, or experience; so it is clear from our passage that the necessity of special laws is a function of what Kant calls 'the order of nature'. I shall discuss the logical status of this conception presently. First, we must see how the distinction referred to bears on what I have called the 'weak interpretation' of the range of the principle of causality. I have argued that the general principle cannot be intended to furnish a justification for the assumption even of lawlikeness in general (let alone the existence of special laws), regarded as holding *between* the phenomena, but that instead causality is no more than a logical ingredient contained *within* each of these phenomena regarded as phenomenal objects of experience. Now the *Critique of Judgement* supposes this interpretation rather more clearly, pointing in a direction reverse to that indicated by the *Prolegomena.* Consider the following passage:

---

[1] *Critique of Judgement*, trans. J. H. Bernard, London, 1892, p. 24

[2] Op. cit. p. 22. Note well, that this 'unity of experience' must be strictly distinguished from the transcendental unity of the 'analytic of principles' discussed in the first *Critique*. That the 'necessity' attaching to natural laws is not 'intuitive' but 'opaque to the intellect' has recently been argued by W. Kneale, *Probability of Induction*, pp. 80 f., 89, 92, 97, 256. However, for Kant, as we shall see, this necessity has no 'ontological significance', but is a function of the spontaneous systematic activity of 'reason'.

[3] Op. cit. p. 24

> The understanding says that every change has its cause. . . . For nature in general (as an object of possible experience) that law is cognized as absolutely necessary.—But now the objects of empirical cognition are determined in many other ways than by that formal time-condition. . . . Hence specifically different natures can be causes in an infinite variety of ways, as well as in virtue of what they have in common as belonging to nature in general ; and each of these modes must (*in accordance with* the concept of a cause in general) have its rule, which is a law and therefore brings necessity with it, although we do not at all comprehend this necessity. . . .[1]

I think this passage requires no further detailed comment. It is quite clear from it that special laws are conceived and formulated in causal terms, ' in accordance with the concept of a cause in general '; but whilst that concept entails a necessitarian logic, there is no question of the general principle of causality being a presuppositional *support* for the special laws. (The phrase ' in accordance with ' should not be construed in that way.) We may sum up the situation perhaps as follows: The Second Analogy does not show that nature is lawlike, but only that the *concept* of law is built into our notion of each objective element of nature. On the other hand, the concept finds its concrete *application* in the formulation of the special laws of nature, which thereby are endowed with a corresponding necessitarian logic. The lawlikeness of the Second Analogy is no more than a *pointer* towards the possibility of the lawlikeness of nature, in the sense of law here understood; and it is this possibility which the Kantian ' reason ' seeks to realise, as the fulfilment of a self-imposed task.

6

Kant frequently harps almost with pleasure on the contingent or ' accidental ' nature of the laws of empirical science.[2] Whatever necessity attaches to them can hence not be ontologically grounded; it depends for substance on the integration of these laws within some scientific system, through which they may be said to exhibit an ' order of nature '.[1] In the first *Critique*, two typical examples of such a 'systematic unity of empirical laws' are Newtonian gravitational theory, regarded as an axiomatic system, and eighteenth-century chemical theory which 'reduces' the multitude of chemical phenomena to three 'elementary principles'.[2] A third paradigm which receives a more detailed treatment in the third *Critique*, are the classificatory systems of botany and zoology.[3]

---

[1] *Critique of Judgement,* op. cit. pp. 21-2 (my italics.)

[2] *Critique of Judgement,* op. cit. pp. 23, 24, 26, 310, etc.

Now the important point about Kant's conception of the 'order of nature' which the construction of such systems implies is that this order is never as such given or presented as a 'constitutive' whole. The reason for this is that there can be no such order outside the systematic and indeed creative constructions of science; these constructions are not 'reflections' of an order which we find, but are the result of a unitary approach which is (in so far as we demand a concept of order at all) 'necessarily presupposed and assumed'.[4] It is in this sense of being 'presupposed', that (what Kant calls) 'reason in its hypothetical employment'[5] proceeds in an '*a priori*' manner, and indeed '*transcendentally*', though only 'in a subjective aspect'.[6] Reason (or judgment) proceed here subjectively only, because the systematic interconnecting of laws does not give us an 'object'; it does not 'prescribe to nature,' but only to reason or judgment itself 'a law for its reflection upon nature'.[7] So without this activity of reason there would be no systematic connection, and hence no sense in ascribing to the special contingent rules of nature any lawlike necessity. But the unity of such a system is not 'given' to us; in the first *Critique* it is described as only an 'aim', or 'projected unity'.[8] The 'principle of such systematic unity', Kant elsewhere calls 'a merely regulative principle and maxim'.[9]

Now this 'regulative approach' is not just a 'psychological matter'. For one thing, we do not proceed in the construction of our systems arbitrarily, but through the introduction of more specialised regulative maxims of judgment (as Kant calls them). Instances of such maxims are ' Nature takes the shortest way ' (the law of economy); ' nature makes no leaps ' (the law of continuity); ' Occam's Razor ', and so forth.[1] Again, at a more general level, as the first *Critique* explains, we employ certain general principles such as those of ' homogeneity, specification and continuity of forms '.[2] Now all such principles or maxims are (according to Kant) to be regarded as neither true nor false; they are not descriptive but prescriptive.

---

[1] That this systemic aspect of laws is what makes certain scientific generalisations lawlike has recently been argued again. Cf. R. B. Braithwaite, *Scientific Explanation*, chap. 9, pp. 300–303 ; E. Nagel, *The Structure of Science*, chap. 4, sect. III. 2, pp. 59–62 ; M. Bunge, *The Myth of Simplicity*, chaps. 10, 2.3, p. 163 ; also p. 174. For Kant this is, however, no merely formal or psychological matter, as will emerge in the sequel.

[2] A662 = B690 ; A652 = B680

[3] Op. cit. p. 24 and *passim*

[4] Op. cit. p. 22

[5] The same notion appears in the third *Critique* as 'reflective power of judgment'.

[6] Op. cit. pp. 23, 25

[7] Op. cit. p. 25

[8] A647 = B675

[9] A680 = B708

> If we propose to set forth the origin of these fundamental propositions and try it by the psychological method, we violate their sense. For they do not tell us what happens, i.e. by what rule our cognitive powers actually operate, and how we judge, but how we ought to judge; and this logical objective necessity does not emerge if the principles are merely empirical.[3]

The notion of ' objectivity ' at which Kant here hints emerges more clearly in the following elucidations. At A647=B675 Kant points out that the existence of verifiable consequences can never ' *prove* ' the universality of an explanatory hypothesis. The hypothetical employment of reason is not ' constitutive ' but only ' regulative '. In bringing systematic unity to the multiplicity of the empirical laws, we can at best hope for the hypothesis to ' approach universality ' indefinitely. The prime purpose of the systematic development of hypotheses is to serve as a testing ground for their truth. So whilst systematisation does not provide proof, it is the only possible method for testing hypotheses.

However, the notion of 'regulativeness ' does not involve mere systematisation. Reason, when employed hypothetically, also introduces a large number of theoretical, simplifying, supplementary conceptions, such as those of a ' pure earth, pure water, pure air '; it is by means of such concepts that simplification first becomes possible.

---

[1] *Critique of Judgement*, op. cit. pp. 20-21

[2] Cf. A658 to A662 (B686 to B690); *Critique of Judgement*, op. cit. p. 24

[3] *Critique of Judgement*, op. cit. p. 21. The importance of regulative principles for the very possibility of an empirical science, Kant once expresses by saying that this possibility constitutes a veritable ' transcendental deduction of all ideas of speculative reason . . . as regulative principles of the systematic unity of the manifold of empirical knowledge in general ' (A671=B699). The ' idea ' here referred to is that of a divine archetype, a ' supremely wise cause ', regarded as having ' the reality of a schema—the schema of the regulative principle of the systematic unity of all knowledge of nature '. It is in this sense, and this sense alone, that we are permitted to ' hypostatize ' the ' objective reality ' of the ' theological idea ' (or rather ' ideal ') of reason. The relevance of Kant's whole treatment of science for his theology cannot be pursued further in this place.

Now such concepts are certainly not ' derived from nature; on the contrary we interrogate nature in accordance with ' them; we ' require them ' for the study of natural phenomena and their causes.[1] So not only does reason involve an aspect of ' creative anticipation ' of nature; but without this feature, a proper testing of both explanatory hypotheses as well as the derivative empirical laws would not be forthcoming. Serious testing involves the prior formation of theories.

Now although the construction of scientific systems does not decide about the truth of their elements, they enable the scientists to extend the range of his researches in any number of possible directions. Such extensions do not predetermine what will and what will not be the case, i.e. empirical *truth*-values; but success in the construction of theories gives evidence of the *correctness* of the method, which consists in forever breaking new ground, and thus is both advanced and confirmed.[2] The new ground, the new insights, due to the introduction of new points of view *via* new theoretical concepts afford the necessary opportunities for the testing of the empirical laws; and it is in this sense that the resulting (though only ' projected ') ' order of nature ' becomes an ' objective ' feature.

Now we have seen[3] that the element of necessity which according to Kant is one of the logical features of empirical laws is stipulated as a condition of there being an ' order of nature '. But this concept itself is now seen to possess a regulative, and not ' constitutive ' significance, although at the same time possessing the ' objective ' power of making a scientific approach possible, and thus preventing these empirical laws from being given in a purely accidental manner. It follows that the necessity of laws must itself be regarded as a pure function of the regulative employment of reason. The ground of this necessity can therefore not be that involved in the argument of the analogies of experience. On the other hand, just as without the necessitarian concepts of the analogies there would be no ' unity of experience ', and indeed, no phenomenal object, so without the regulative employment of reason, and its consequent regulative necessity (through its methodological maxims) there would be no empirical science and consequently no ' order of nature '; the difference between the two levels being that empirical objects can be said to be given, whereas the ' unity of science ' is only ' aimed at '; its realisation is only ' set as a task '.[4]

---

[1] A646=B674 [2] A680=B708 [3] Above, Section 5

[4] A508=B536. However, we must remember that ' objects ' are ' given ' only subject to the categories. And just as ' objects ' require the ' spontaneous ' aid of the

Whilst it is true that without the ' regulative employment of reason ' the study of nature could not even begin, the pursuit of its tasks cannot vouchsafe its completion; how much less can the analytical principles, such as that of causation, guarantee the lawlikeness of nature.

7

To illustrate the relation between the two levels further, we must study also the place of the causal principle in the third *Critique* (par. 70), where we get a discussion of the status of determinism, contrasted with the principle of teleological causation. For mere empirical judgment, Kant tells us, we require no more than the general analogies and those ' particular determinations ' met with in experience, i.e. concrete happenings, such as a ship's moving down-stream. But as regards the particular laws of science, we require an additional ' guiding thread ',[1] if we are to ' hope for ' any systematic explanations. To this end, we start with two ' maxims ', one of which leads to the mechanistic approach, the other to the teleological. Now one should note carefully how the principle of mechanistic determinism is here introduced. The maxim is, Kant tells us, ' *a priori* suggested by the mere understanding.' So the principle is not claimed to be either equivalent to, nor deducible from, the principle of causality; it is only ' *suggested* ' by the latter. It follows that both maxims in question are once more purely ' regulative principles '.[2] It is indeed very important for Kant to be able to regard the principle of determinism as merely regulative, so that he can say ' judge nature in accordance with mechanical laws ', and not: ' all things in nature are possible only in accordance with mechanical laws.' It is the former locution alone that allows him to use alternatively either the deterministic or the teleological approach.

---

understanding (through the categories), so the further pressure of ' reason ' drives these objects into systematic relations (in the theoretical unities of science), guided by regulative maxims, whose ' archetype ' is eventually located in the image of a God. The critical suspension of the independence of ' objects ' by the understanding (the restraint of reason) is to make possible the creative development of ' nature ' in these systematic constructions, and ultimately, the archetypes of religious experience. It was no empty gesture when Kant wrote in the ' Preface to the Second Edition ' of the *Critique of Pure Reason* : ' I have therefore found it necessary to deny *knowledge*, *in order to* make room for *faith* ' (Bxxx). For Kant's development of this theme, cf. A696=B724 f.

[1] Cf. op. cit. p. 294, for this expression ; cf. also p. 25

[2] Ibid.

The notion of purpose pervades of course (as is well known) the whole of the third *Critique*, though it occurs already in the first.[1] For the argument of this paper we have however been able to ignore this aspect. In its *general* signification, purpose is introduced by treating the systematic unity of empirical laws 'on analogy' with something so designed as to be adapted to an understanding which regards the world as systematically organised in accordance with certain 'purposive principles', namely those very regulative maxims of judgment, referred to already (e.g. laws of economy, continuity, etc.). Now Kant thought that he possessed indeed a paradigm case of such an approach in physics, consisting in the 'unification of different rules apparently quite independent of one another in a single principle',[2]—almost certainly a reference to Maupertuis' 'Principle of Least Action'.[3]

To Maupertuis the discovery of a principle involving the notion of economy had seemed to offer a 'proof' for the existence of God. According to Kant, a teleological principle does not of course point to a designer; nevertheless the teleological standpoint, regarded from a 'regulative point of view' is useful in 'supplementing the insufficiency of the mechanical laws', when we turn to the attempted systematisation of the special laws of mechanics, statics and optics (as in the work of Maupertuis). In other words, the regulative approach, under its teleological guise could on occasion yield *specific theoretical principles*, and in this way prove genuinely creative. It is in this further sense, that the regulative maxims determine what ought to be, and not what is; and it follows that in this sense the laws of nature form a system, an order, and acquire a necessitarian logic.

It is quite clear then, that to the extent to which Kant shifts the emphasis upon the regulative or prescriptive power of scientific constructiveness ('reason'), the lawlikeness of laws must be made dependent on reason and not the understanding. We may conclude then that this lawlikeness is a notion which has a multiple interpretation in the Kantian corpus: (1) as a transcendental presupposition of self-contained contingent objects (or more properly: objective states of affairs); (2) as a regulative injection of the concept of cause, in the attempt to fit causes to the actual changes of nature; (3) as a more determinate specification, under which it appears as the principle of mechanistic determinism; (4) as the idea of the systematic unity of nature, under the guise of theories, constructed through the guidance of certain regulative maxims. In particular, when Kant says that the analogies give us causality, but that special laws require experiential specification for the discovery of actual causal laws, we must not take it that the general principle provides a justificational basis for such laws, or for their ' lawlikeness in general '; that is provided by the regulative function of reason.

---

[1] Cf. A686=B715 ff; A697=B725 f

[2] *Critique of Judgement*, par. 68, op. cit. p. 289

[3] Kant commentators have not usually noticed the relevance of this paragraph nor the implied reference to Maupertuis, who whilst not mentioned here by name, had come early under Kant's notice, through an essay published in 1744, bearing a title almost identical with the passage just cited; cf. also Kant, *On the Only Possible Demonstration for the Existence of God*, ed. Hart., ii, pp. 141-143.

8

Of course, like all interpretations of Kant, there are passages which cannot be made to fit too neatly to the point of view which has here been developed. In particular, there is one special case which seems openly to contradict it, and that is the treatment of the primary axioms of the Newtonian system, e.g. the laws of conservation of mass, of motion, and of momentum, especially as developed in the *Metaphysical Principles of Natural Science*.[1] Consider the law of inertia. Kant's ' proof ' is as follows. First he argues that from general philosophy we take the principle of causality, here formulated as the principle that every change has a cause. We then add the following specifications, partly empirical, partly conceptual: in mechanics, change means change of velocity; cause means force, and in particular: external force. Finally, all matter, quâ matter, is ' inert '. Given these ingredients, according to Kant we then get the law of inertia. Now here, rightly or wrongly, the law of causality seems to be intended to *support* the law of inertia. And the reference to ' general philosophy ' seems clearly intended for the argument in the *Critique*. This means, the law of causality seems now applied, not as a presupposition of our cognition of change, but of a *law* of change.

There are two positions we can take: (1) Admit that there is a straight contradiction, or cleavage, running through Kant's thinking. (2) Insist that the present case must be treated on the lines I have suggested. In other words, the law of causality, though analytically a general presupposition, is here *applied to*, though not regarded as *supporting*, a special case. It is brought in, in order to help us ' to search for and to formulate '[1] specific laws of nature. Whilst this would not be in line with the bold claims which Kant is often believed to have advanced on behalf of Newtonian science, it would offer a more defensible approach, and one which is more consistent with the general aspects of Kantian philosophy.[2]

---

[1] Part III, Theorem 3. Hart. ed., iv, pp. 439-440

University of Cambridge

---

[1] A545=B573 ; cf. above, p. 198

[2] That Kant in this proof of the law of inertia goes in for a certain amount of special pleading, emerges also from the following consideration. As the example of the ship moving down-stream implies, no restrictions are at that level of Kant's argument applied to the kind of change intended. However, a few pages later (A207=B252) Kant suddenly remarks in a footnote : 'It should be carefully noted that I speak not of the alteration of certain relations in general, but of alteration of state. Thus, when a body moves uniformly, it does not in any way alter its state (of motion) ; that occurs only when its motion increases or diminishes.' This restriction is however practically equivalent to the law of inertia.

CONNECTIVITY, CHANCE, AND IGNORANCE*[1]

D. H. MELLOR

> ' Ignorance is like a delicate exotic fruit ; touch it, and the bloom is gone '
> —Lady Bracknell in Oscar Wilde's *The Importance of Being Earnest*, Act I

I

CONSIDER the following instance, p_1, of what I shall call a ' chance proposition ':

' The chance of an unbiassed coin falling heads is 1/2.'

Is p_1 an empirical assertion or not? There are two obvious possibilities:

1. ' Bias ' is defined in terms of physical properties of the coin without reference to the chance of it falling heads. Then p_1 is a natural law, if true. Confronted by a counter-example, we should reject p_1, if not with equanimity, at least without a sense of conceptual outrage.
2. p_1 is analytic, since ' bias ' is defined in terms of the chance of falling heads. Any apparent counter-example would be rejected as *ipso facto* biassed.

Neither of these alternatives is appealing. p_1 does not seem to be analytic in the trivial definitional sense of (2), yet it would surely not be as readily abandoned as (1) suggests. The first aim of this paper is to display p_1 as an instance of a general regulative principle in science, namely that of connectivity. The principle justifies the correct conclusions of classical probability arguments, while showing their limitations. In particular, it does not rely on arguments from ignorance, and so evades the notorious paradoxes. On the other hand, it does not need to retreat into relative frequency, subjectivism, logical probability or the synthetic *a priori*.

In Section 2, the notion that certain chance propositions express physical properties is set out and briefly defended. In 3, the principle of connectivity is introduced and, in 4, applied to these chance propositions. In 5 and 6, the connectivity argument is contrasted with classical arguments from ignorance, and, in 7 and 8, an influential variant of the classical view is critically examined.

* Received 1. iv. 65

[1] I am especially indebted to Dr M. B. Hesse and Dr I. Hacking for pointing out substantial faults in an earlier version of this paper. They are not responsible for any that remain in this version.

2

Consider the following chance proposition:

'The chance of this coin falling heads is 1/2.'

I follow Hacking [7], Popper [14], Braithwaite ([1], p. 187), and Peirce ([12], p. 169) in supposing this proposition, p_2, to state a (dispositional) physical property. Further, it is not just a property of the coin, but of the way it is tossed and the surroundings in which it is tossed. Call all this the 'set-up' ([7], p. 3), ([14], p. 67). Then p_2 expresses a property of the set-up, which we will expect to be connected with its other properties.[1] For example, if the coin is magnetised, the chance of it falling heads may be affected by the presence of a magnetic field. A rise in temperature may warp the coin. A change in the tossing machine may affect the chance of the coin falling heads.

I am not concerned in detail with how chance propositions like p_2 are tested, and what exactly their relation is to propositions about relative frequencies in finite sequences of trials. I assume that there are agreed criteria for accepting or rejecting p_2, if only provisionally. ([1], chap. 6). But it may be thought that chance is sufficiently peculiar in its logic of measurement and hence empirical meaning to disqualify it from being classed as a physical property in the same sense as temperature, volume, electric current, pressure and the like. It is worth a digression to show that this is not so.

No doubt, only chance propositions derived from p_2 of the form:

'The chance of this coin falling heads lies in the interval (a, b)',

where $0<a<\frac{1}{2}<b<1$, can be directly tested, and not p_2 itself. Moreover, there always remains a finite chance that, on any given test criterion, a true proposition of this form will be rejected or a false one accepted ([11], chap. 1, pt. 3).

But, as Chwistek has observed ([5], p. 256), all this is just as true of propositions assigning particular values of temperature, volume, field strength, or any other quantity measurable on a continuous interval scale. Take temperature as a typically respectable dispositional physical property, and consider the temperature proposition:

'The temperature of this coin is 21°C.'

Again, all that can be directly tested with thermometers of varying precision are derived propositions of the form

'The temperature of this coin lies in the interval (a,b) °C', where $-273.2<a<21<b$. Moreover, there always remains a finite chance that, on any given test criterion (i.e. method of reading a given thermometer), a true proposition of this form will be rejected or a false one accepted.

---

[1] For brevity, I shall continue to use such terms as 'unbiassed coin', 'true die', but these are to be understood as referring to properties of the whole set-up.

To this parallel, it may be objected that, for a chance proposition such as p_2, the chances of error are calculable from, and intrinsic to the meaning of, the original proposition, and that this is not true of a temperature proposition. But this is not so. Fluctuating molecular velocities limit the precision of any temperature measurement. It follows from statistical mechanics, that there is a finite chance of a particular measured value deviating by any given finite amount from the 'true' mean value. This is certainly intrinsic to the (theoretical) meaning of temperature, and the chances are calculable from it. The same, of course, is true of lengths, pressures, electric currents, and the magnitudes of fields, electro-magnetic, gravitational, connected with particles in indeterminate motion.

Conversely, there are *incidental* sources of error also in tests of chance propositions. It is as possible to make a mistake in counting the number of heads in a sequence of tosses as it is to misread a thermometer. It is as possible for a particular sequence to misrepresent the chance (if, say, the coin is warped during that sequence by a temperature change) as for a particular surface reading to misrepresent the temperature (if, say, temperature gradients are set up in the coin by chemical action). Intrinsic sources of error are often more noticeable in tests of chance propositions, but this is merely a difference of degree, contingent on the customary precision with which different quantities are measured.

It may yet be urged that, in giving the (empirical) meaning of a chance proposition such as p_2 in terms of its conditions of confirmation, yet another chance proposition has been invoked, namely that of making an error of measurement. In contrast, it is not necessary to invoke one temperature proposition in giving the empirical meaning of another temperature proposition. Consequently, the meaning of chance propositions is held to be peculiarly obscure. But it has been shown that the empirical meaning of a temperature proposition also involves chance propositions about errors of measurement. Why should any obscurity in their meaning be held not to infect that of the temperature proposition? In any case, I take the empirical meaning of chance propositions such as p_2 to be sufficiently given by Braithwaite's hierarchy of rejection rules ([1], chap. 6), in which successive propositions state the chance of error in applying the preceding rejection rule. 'It is true that a rejection, or any series of rejections, never "yields certainty" in the sense that all these rejections may be cancelled by a further test' ([1], p. 163), but this, as we have seen, is just as true of temperature as of chance propositions.

That chance propositions differ in meaning from temperature propositions is as evident as that the latter differ in meaning from propositions about length or chemical composition. All I have been concerned to show is that there is not, as might appear at first sight, a peculiarity in the logic of measurement of chance which should lead us to deny that it *can* be a dispositional physical property of a set-up in the same sense as temperature and the like.

Now it is indeed obvious that *some* chance propositions do *not* state a physical property. This fact may have contributed to a false impression that *no* chance proposition does so. I think this confusion of usage is sufficiently important to be worth resolving, conformably to the recent literature, by reserving 'chance' for statements about the physical property. Section 6 analyses an example of a chance proposition which appears not to state a physical property. The aim of the analysis is to support by illustration the proposed linguistic expression of the familiar distinction between the physical property of chance and the degree of confirmation, or inductive probability, of a hypothesis (see, e.g. Carnap [4]). Meanwhile, it remains to show just what of relevance is implied by saying that a proposition, chance or other, states a physical property of a set-up or, in general, of any physical system. This is the object of the next section.

3

The principle of connectivity has recently been expounded by Schlesinger ([16], chap. 3), drawing on the works of Maxwell [17], Campbell ([3], chap. 3) and Bridgman ([2], chap. 3, 'The Causality Concept'). Briefly, the principle asserts that 'two physical systems never differ in a single aspect' ([16], p. 73) or property. If they did, the difference could not be explained by being correlated with any other. This is so whatever view—positivist, instrumentalist, or realist—one takes of theoretical explanation. To construct—or employ, or discover—a theoretical entity to account for just one isolated difference is to restate the difference rather than to explain it. That a theoretical entity detectable in just one way is unacceptably *ad hoc* is a corollary of connectivity that is probably better recognised than the principle itself.

As Schlesinger observes, the principle of connectivity does not make a straightforwardly empirical assertion, since 'we are not given a clear, detailed and generally applicable definition of what constitutes a single aspect' ([16], p. 88). Any proposed counter-example could apparently be dealt with in one of two ways:

1. One could refuse to recognise the observed difference as exhibiting a real property or physical state at all. Thus Schlesinger: 'In general, according to the principle of connectivity, if two physical systems have different properties or the same property to a different degree, this is bound to manifest itself in more than one way' ([16], p. 78). Turner similarly expresses Maxwell's view that '. . . if a quantity is connected to other effects which are independently defined then it is a physical state; if not, then it is a mere scientific concept' ([17], p. 231).

2. One could say that the observed difference was not really in a 'single aspect', that 'we could always break up the ostensibly single property into more elementary ones' ([16], p. 90). Thus Poincaré's example[1] of a substance differing from phosphorus only in its melting point could be dealt with by saying that 'melting point alone may be exhibited as a complex property; it is the property of becoming liquefied at a certain temperature but also the property of discontinously assuming a different specific heat at the very same temperature' ([16], p. 91).

Now this latter example is unconvincing just to the extent that the connections between the different properties which characterise a melting point are well-established empirically, deeply embedded in our structure of physical laws and theories, and central to the conceptual scheme (of solids and liquids) which these provide. To this extent, the concept 'melting point' embodying these connections *is* regarded as a real single physical property, and hence to be explained by being connected through the law structure with other such properties. To this extent, therefore, one would refuse to evade the application of connectivity to it, and insist that any observed difference in melting points must be connected with some other physical difference. Only after the most strenuous investigation had failed to disclose such another difference would one be prepared to reconstruct the conceptual scheme and ' deny that a pure substance had always a constant melting point ' ([3], p. 76), or that melting point is a single physical property.

---

[1] 'When I say : Phosphorus melts at 44°, I mean by that : all bodies possessing such or such a property (to wit, all the properties of phosphorus, saving fusing-point) fuse at 44°. . . . Doubtless the law may be found to be false. Then we shall read in the treatises on chemistry : "There are two bodies which chemists long confounded under the name of phosphorus ; these two bodies differ only by their points of fusion."' ([13] 'The Value of Science', pt. 3, chap. 10, sect. 4, p. 333). See also ([13] 'Science and Method', bk. 2, chap. 4, sect. 12, p. 470).

But one way or another, the principle of connectivity would be preserved. It appears then as a regulative criterion of adequacy for the scientific characterisation and explanation of a system. Consequently, one need not suppose every minute detail of the system to be accounted for, but certainly all the physical states, the properties linked by the structure or network of accepted laws which apply to it. In general, to point to some feature of a system as representing a real physical (i.e. scientifically explainable) property is to require, among other things, that connectivity be applied to it.

4

The relevance of the above discussion of connectivity is to make clear what is implied in claiming that the chance proposition p_2 states a physical property of a set-up. It is implied that the principle of connectivity is applicable and hence that two such set-ups, S_1 and S_2, cannot have every other physical property the same, and yet differ in the chance of the coin falling heads.

Now those physical differences which serve to label one side of the coin ' heads ' and the other ' tails ' are certainly not in the set of properties by which a difference of chance is to be explained. If they were, if labelling the sides of a coin generally biassed it, the explanation of the bias would be trivial, and a coin would be unacceptable as part of a chance set-up. In short, these differences are not taken as physical properties of the set-up.

It follows that the application of connectivity is unaffected by relabelling the sides of the coin in, say, S_2. The condition for qualitative identity of S_1 and S_2 would then be that every true proposition ascribing a physical property to S_1 should become true of S_2 when ' tails ' is substituted for ' heads ' (and vice versa) whenever the latter occurs in the sentence expressing the proposition. For example, if the coins are magnetised and the N-pole in S_1 is heads, in S_2 it must be tails. If this condition is satisfied, connectivity requires that the chance of S_1 falling heads shall be the same as that of S_2 falling tails.

Now consider what follows when the chances of S_2 falling heads and tails are unequal. Denote the chance of S_1 falling heads '$P_1(h)$', and the chance of it falling tails '$P_1(t)$'. Denote the set of those other physical properties of S_1 which are peculiar to heads 'h_1', and the set of those peculiar to tails 't_1'. Similarly for S_2. Then the two applications of connectivity above lead to the assertions

$$h_1 \equiv h_2 . \supset . P_1(h) = P_2(h) \tag{1}$$

and

$$h_1 \equiv t_2 . \supset . P_1(h) = P_2(t) \tag{2}$$

Now suppose that (1) is exemplified by S_1 and S_2, i.e. that

$$h_1 \equiv h_2 \tag{3}$$

Hence, from (1) and (3),

$$P_1(h) = P_2(h) \tag{4}$$

Suppose, moreover, that

$$P_2(h) \neq P_2(t) \tag{5}$$

(e.g. $P_1(h) = P_2(h) = 0.7$).
Hence, from (4) and (5),

$$P_1(h) \neq P_2(t) \tag{6}$$

Hence, from (2) and (6),

$$h_1 \not\equiv t_2 \tag{7}$$

But now, from (3) and (7),

$$h_2 \not\equiv t_2$$

Thus, for any coin-tossing set-up, S_2, given the existence of a set-up, S_1, related to it by (3), connectivity requires that a difference between $P_2(h)$ and $P_2(t)$ shall be accompanied by a difference between h_2 and t_2. But the existence of a set-up, S_1, is entailed by the existence of S_2, since the latter is itself such a set-up, whose self-identity entails the truth of (3).

Hence, for any coin-tossing set-up, S, connectivity requires a difference in the chances of S falling heads and tails to be connected with some other difference between those physical properties of the set-up peculiar to heads and those peculiar to tails. Conversely, if the coin is unbiassed in the sense that there is no such other difference, then connectivity asserts that there is no difference either between the chance of falling heads and the chance of falling tails. If the coin must fall either heads or tails, then we arrive again at proposition p_1 (Sect. 1). But the sense in which p_1 is stronger than a straightforward empirical assertion while not being trivially analytic is now clear. By regarding such chances as being properties of the set-up, we require differences in them to be explained by correlation with other physical differences. p_1 merely states one instance of this regulative principle.

5

It is necessary to elaborate on the conclusions of Section 4, to contrast them with the results of a classical analysis. First, the chances asserted to be equal need not be exhaustive. If there is a 0.02 chance of an unbiassed coin landing on its edge, connectivity gives its chance of falling heads as 0.49. Second, the argument does *not* rest upon ignorance. In other words, there is no question of having to carve up the entire range of possibilities into 'a certain number of cases equally possible, that is to say, to such as we may be equally undecided about in regard to their existence' ([9], p. 6). The important point is the redundancy and invalidity of arguments from ignorance, and the rest of the paper is devoted to illustrating this.

It has been assumed in discussing p_1 that the coin is known to be unbiassed in all relevant properties, and the relevant properties are those connected with the chance of falling heads or tails by the currently accepted law-structure. It should perhaps be made clearer that this invoking of relevance does not after all render p_1 trivially definitional.

Inferences from propositions about other physical properties to chance propositions are deductive, given the appropriate laws. But the laws themselves—such as those connecting the magnetic properties of a coin with the chance of it falling heads in a set-up of a certain kind—can only be supported by empirically observed correlations of chances with other properties. For any specified coin, valid inference from possession of other properties, Q, R, say, to the chance of falling heads is contingent on the truth of the particular laws connecting the chance to Q, R; the truth of such laws can, of course, never be known with certainty.

Consequently, one may be wrong in thinking that a coin is unbiassed, either because of some mistake in the measurement or description of the set-up (e.g. wrongly ascribing possession of Q or R to it), or because a law relating Q or R to the chance has been assumed which is in fact false. Moreover, even if all the known laws are true, one may be unaware of the sufficient conditions for their truth; i.e. because the completeness of the law structure can never be established, one may be ignorant of some other relevant properties. Thus the these circumstances, not to know which way the coin is biassed is just not to know how its sides are labelled. It is, after all, irrelevant that the sides of a coin are usually labelled when it is made and its physical properties determined. This need not be so: a coin may become biassed, e.g. by bending, after it is labelled, or the bias may arise from the structure of the tossing machine. Conversely, the labelling may be reversed by an arbitrary reinterpretation of the surface features of the coin.

In short, there is a good deal of empirical evidence that labelling and bias manufacturing processes involved in producing a coin-tossing set-up are quite independent. This is what justifies the assumption made in Section 4 that labelling a coin does not bias it. It also justifies the converse assumption that biassing a coin does not at all determine the way it is labelled. This is the assumption implicit in the described situation in which each side has an equal chance of being labelled ' heads ', although the coin is known to be biassed.

In this light we may now attempt a more rigorous analysis to derive p_4 for the biassed coin. In the first place, the bias must be used to distinguish the sides independently of the labelling. To talk of the chances of two indistinguishable, unlabelled sides receiving a certain label would be incoherent. In one of the intermediate cases referred to above, where it is not known whether the coin is biassed or not, we may proceed as follows. If the coin is unbiassed, p_4 is derivable as above. If the coin is biassed, p_4 is derivable as shown below. Thus if p_4 can be derived for a coin known to be biassed, it can be derived for a coin merely suspected of bias. The necessary restriction of the ensuing argument to coins known to be biassed is therefore of no consequence.

Denote the side of the biassed coin that has the greater chance of landing uppermost in a toss ' the likely side ', and the other side ' the unlikely side '. Now consider the following chance proposition, p_5:

' The likely side of this coin has a chance of being labelled " heads " of $\frac{1}{2}$ '.

This proposition states a physical property of a coin-labelling set-up. Then the empirical assumptions made above come to this: that those features of a coin which distinguish the likely from the unlikely side, i.e. which *are* properties connected to the chance in a coin-tossing set-up, are *not* connected properties in the coin-labelling set-up. Since, *ex hypothesi*, these are the only features, other than those introduced by labelling, which distinguish the two sides, *no* physical property connected to the chance of labelling 'heads' is peculiar to the likely side. Hence, by a connectivity argument parallel to that of section 4, it follows that the likely and unlikely sides have equal chances of being labelled 'heads'. Since one or other side must be so labelled, we arrive at p_5.

The general principle, P, may now be invoked on the following interpretation. A trial T of kind K is the labelling of the coin by a specified labelling set-up. The outcome E is that the likely side is labelled 'heads'. Then p_5 states that the chance of E on trials of kind K is $\frac{1}{2}$. Hence we conclude from P that the probability, $c(h_2)$, of the hypothesis, h_2:

'The likely side of this coin will be labelled "heads" on trial T', is $\frac{1}{2}$. Denoting the corresponding hypothesis about the unlikely side of the coin 'h_3', we then have

$$c(h_2) = c(h_3) = \tfrac{1}{2} \tag{8}$$

The derivation of p_4 from equations (8) follows at once. Let the chance that, in the coin-tossing set-up, the coin will fall with the likely side up be q $(q>\frac{1}{2})$. Hence, by P, the probability, $c(h_4)$, of the hypothesis, h_4, that this will occur on a given toss is q. If the coin must fall one or other side up, the probability, $c(h_5)$, of the hypothesis, h_5, that this will be the *un*likely side is $1-q$. Thus, by an elementary theorem of probability (e.g. Feller [6], p. 106), the probability, $c(h_1)$, of h_1 is given by

$$\begin{aligned} c(h_1) &= c(h_2)\,.\,c(h_4)+c(h_3)\,.\,c(h_5) \\ &= \tfrac{1}{2}\,.\,q+\tfrac{1}{2}(1-q) \\ &= \tfrac{1}{2}. \end{aligned} \tag{9}$$

Equation (9) states the required proposition p_4.

Thus, even in the case of the biassed coin, the correct conclusion of the classical argument follows, not from ignorance, but from knowledge of those characteristics of coin-tossing and coin-labelling set-ups which make the former suitable as statistical devices for games of chance. We may thus identify the proposition p_3 with the proposition p_5 and conclude that it is, after all, a chance proposition stating a physical property, not of a coin-tossing, but of a coin-labelling, set-up. By arguing from ignorance for both the biassed and unbiassed coins, the classical view fails to exhibit the radical difference between the chance propositions which are required in each case to sustain the same probability proposition.

7

It is obvious that the connectivity arguments applied in section 4 to coin-tossing set-ups similarly justify the ascription of equal chances to throwing each face of a true die. Since the argument does not rest on ignorance, the well-known paradoxes are easily avoided. For example, ' throwing a 6 ' and ' not throwing a 6 ' are known not to be equiprobable alternatives, and similarly with ' throwing a 5 ' or any other number.

In general, the use of connectivity to justify ascribing equal chances is restricted to systems whose other properties are known to be symmetrical in the relevant respects. In games of chance, of course, to which classical probability was originally, and is appropriately, applied, this symmetry is deliberately contrived. Even so, in any particular case, positive empirical evidence must support any claim to knowledge of the symmetry.

It is even more necessary to insist on this in investigating uncontrived natural phenomena, especially since Kneale ([8], pt. 3) has made an influential attempt to extend an essentially classical view to such cases. Since the success of his attempt would be fatal to the thesis of this paper, it is worth examining in some detail. Kneale considers the chance that an α-thing is also β, and tries to resolve the difficulty that there may be more than one equally fundamental measure of a range of α-ness, leading to conflicting values of the chance. That he regards the success of this resolution as crucial is shown by his assertion that ' to admit two different ways of measuring the range would be to abandon all hopes of formulating an objectivist theory of probability ' ([8], p. 183).

Part of Kneale's argument is perfectly acceptable, since he is prepared to use empirical information about a set-up to exclude certain measures as inappropriate. For example, he shows how the ' puzzles of Bertrand's paradox disappear when a practical method for selecting a chord at random is specified ' ([8], p. 185). But this is not enough, since an infinite number of non-linear measures of, say, area remain, which only positive empirical information justifies us in excluding. The general point, that to know enough about the set-up to choose the appropriate measure is to know enough to determine directly the chance that the measure is supposed to establish, has been made by one reviewer ([18], p. 23), but the weakness of Kneale's answer to this anticipated objection ([8], p. 188) has not been sufficiently exposed.

Kneale relies on a generalisation of a theorem of Poincaré's ([8], pp. 142–144) ([13] ' Science and Method ', bk. 1, chap. 4, sect. 6, p. 403). Poincaré takes the example of a spinning pointer, *r*, and considers the chance, $f(x)$, that it will travel an angular distance less than x before stopping. He starts from the reasonable assumption that this cumulative probability function, $f(x)$, is a continuous and analytic function of x, which he takes as the measure of the range. Then, for any $f(x)$, if a and b are two *sufficiently close* values of x, it certainly follows that the chance of *r* stopping between a and $(a+b)/2$ differs from the chance of *r* stopping between $(a+b)/2$ and b by less than any stipulated amount. however small. Kneale then argues that if the pointer is spun hard enough, these ' sufficiently close ' values, a and b, may be taken to differ by 2π, so that whatever the overall shape of the probability function, there is an effectively equal chance of the pointer stopping in any equiangular sector of the circle. But this further inference is invalid, and the trouble lies in the words ' sufficiently close '. That a sufficiently small interval (a,b) *exists* is true; that this interval contains the interval of interest (in this case, one of length 2π) does not follow at all, however short the interval of interest may be.

Consequently, the general conclusion which Kneale draws from Poincaré's theorem, and on which his whole theory admittedly depends, is simply false:

> ' . . . if, as seems reasonable, it can be assumed (*a*) that of any two connected variables which are equally fundamental each is a continuous analytic function of the other and (*b*) that by either scale of measurement the range of α-ness is a small part of the configuration space in which it lies, we are entitled to say that metrical relations *within* the range of α-ness will be approximately the same, whether the variation of x or that of y be chosen as a dimension of the configuration space ' ([8], p. 189).

It is not enough to know that ' the range of α-ness is a small part '; it must be a *sufficiently* small part for the degree of approximation required, and to know *this* is already to know the conclusion which it is desired to establish. The same objection holds against his further analogy from map-making:

> ' When two maps of a continent are made according to different projections, regions represented by equal areas in one map may be represented by markedly unequal areas in the other. If, however, we consider only those parts of the same two maps which correspond to a single county, we find that the differences introduced by different

> methods of projection are negligible; parishes which are represented in the one by equal areas are represented also in the other by approximately equal areas' ([8], p. 189).

Again, the problem remains, given the degree of approximation, to show that a particular parish or county is *sufficiently* small to be represented by equal areas on the two projections in question, and this cannot be known *a priori.*

8

The invalidity of the classical argument and the proper application of connectivity may be illustrated with Kneale's own example. Let x be the angular distance travelled by a spinning pointer, and let another measure of the range be

$$y = x + m \,.\, \sin(x), \text{ where } m < 1 \text{ is a constant.}$$

x and y are, on Kneale's criteria, equally fundamental ([5], p. 187), and each is a continuous analytic function of the other. Then, if the cumulative probability function, $f(x)$, is continuous and analytic, so is its counterpart, $g(y)$.

Now suppose $g(y)$ is such that any interval of length

$$y_1 - y_0 = 2\pi = x_1 - x_0$$

is 'sufficiently small', i.e. such that the chance of the pointer stopping in any small sub-interval, (a,b), whose measure, $\triangle y$, is less than 2π, is directly proportional to $\triangle y$. Writing this chance '$\mathrm{P}(a,b)$', we have

$$\mathrm{P}(a,b) = k \,.\, \triangle y, \text{ where } k \text{ is a constant.}$$

In particular, this is true for a very small sub-interval (a,b), such that its measure, δy, is given by

$$\delta y = \frac{dy}{dx} \,.\, \delta x.$$

Hence, $\mathrm{P}(a,b) = k \,.\, \delta y = k \,.\, \frac{dy}{dx} \,.\, \delta x = k(1 + m \,.\, \cos(x)).\delta x.$ Hence the chance that the pointer will stop in a small sector of angular width δx is $k(1 + m \,.\, \cos(x)) \,.\, \delta x$. This chance is, therefore, *not* simply proportional to δx, but varies between $k(1+m) \,.\, \delta x$ at $x = 2n\pi$ and $k \,.\, \delta x$ at $x = (2n+1)\pi$, where $n = 0, 1, 2. \ldots$ Clearly, there is an indefinite number of such measures, y, for different values of m. One or other could well describe the behaviour of a magnetised pointer in a magnetic field, however hard the pointer was spun. In this case, as in every other, all that can justify excluding such a possible measure of the chances is positive evidence that the relevant connected conditions—e.g. presence of a magnetic field—do not obtain.

To say this is to take such a probability measure as stating a physical property of the set-up in the sense given in sections 2-4 above and consequently to insist on the application of connectivity. Then, if there are sectors of the circle between which no other physical property discriminates, connectivity requires that the chances of the pointer stopping do not discriminate between them either. And in a particular case, these sectors may be known to be equiangular.

The analogy with the coin-tossing set-up may be completed by considering the case where the pointer is in a known magnetic field, and it is known that it is magnetised, but not which way. Then one might properly conclude that there was an equal chance of the pointer stopping with its labelled end to one or other side of the line through its axis perpendicular both to the axis and the field. As with the biassed coin of Section 6, this chance proposition states a physical property, not of the pointer-spinning set-up, but of the pointer-labelling set-up.

In all these cases, the desired conclusion can properly be drawn only from sufficient knowledge of symmetry or independence. As with all empirical knowledge, this knowledge is never certain. The claim that it is sufficient is just as needful—no more, no less—for the inference from it to a chance as to any other proposition. This claim is never, as Kneale thinks, demonstrable; if called in question, it can only be supported by collecting more empirical evidence.

Pembroke College,
Cambridge

REFERENCES

1. R. B. Braithwaite, *Scientific Explanation*, Cambridge, 1953
2. P. W. Bridgman, *The Logic of Modern Physics*, New York, 1927
3. N. R. Campbell, *Physics : the Elements*, Cambridge, 1921 ; reprinted as *Foundations of Science*, New York, 1957
4. R. Carnap, 'Statistical and Inductive Probability' in *The Structure of Scientific Thought*, pp. 269-279. (Ed. E. H. Madden.) Boston, 1960
5. L. Chwistek, *The Limits of Science*. (Trans. Brodie and Coleman.) London, 1948
6. W. Feller, *An Introduction to Probability Theory and its Applications* (2nd edn.), vol. 1. New York, 1957
7. I. Hacking, 'On the Foundations of Statistics', *British Journal for the Philosophy of Science*, 15, (57), 1964, 1-26
8. W. C. Kneale, *Probability and Induction*. Oxford, 1949

9. Marquis de Laplace, *Philosophical Essay on Probabilities, 1814* (Trans. Truscott and Emory.) New York, 1951
10. E. Nagel, 'Measurement' in *Philosophy of Science*, pp. 121-140. (Ed. Danto and Morgenbesser). New York, 1960
11. J. Neyman, *Lectures and Conferences on Mathematical Statistics and Probability*, New York, 1952
12. C. S. Peirce, *The Philosophy of Peirce*. (Ed. J. Buchler.) London, 1940
13. H. Poincaré, *The Foundations of Science*. (Trans. G. B. Halsted.) Science Press, Lancaster, Pennsylvania, 1913
14. K. R. Popper, 'The Propensity Interpretation of the Calculus of Probability . . . ', in *Observation and Interpretation*, pp. 65-70. (Ed. S. Körner.) London, 1957
15. B. Russell, *The Principles of Mathematics* (2nd edn.). London, 1937
16. G. Schlesinger, *Method in the Physical Sciences*, London, 1963
17. J. Turner, 'Maxwell on Method of Physical Analogy', *British Journal for the Philosophy of Science*, **6** (23), 1955, 226-238
18. F. L. Will, 'Kneale's Theories of Probability and Induction', *Philosophical Review*, **63**, 1954, 19-42

THE CONCEPT OF PROBABILITY*

NEIL COOPER

THE concept of probability has its origins in everyday thinking and it is only since the seventeenth century that it has acquired a technical and numerical use. Because of its history, two kinds of theory have grown up about the nature of probability, one-concept theories and two-concept theories. Those who hold two-concept theories very often pooh-pooh the idea that there is any connection between the uses, or even suggest that one of the concepts is just uninteresting. Both these positions are held in the following conversation (*Observation and Interpretation*, ed. Körner, Dover, p. 80):

> *Darwin*: We have all been physicists and we tend to think about probability on the basis of relative frequency. . . . But there has got to be a theory of probability which is not a frequency theory, as I can show by citing the case of the horse-race. When I bet and take the odds as four to one on the horse winning the race, I do not mean that I know that this horse is going to run the race 100 times and to win 20 times; I am absolutely certain that it is not going to run 100 times; therefore I have got to have an independent, separate idea of probability. . . .
>
> *Vigier*: I think the case of the horse-race and the case where you have relative frequencies are two separate questions. The word ' probability ' has just introduced confusion. The word has a meaning only in the second case where you have an ensemble and calculate relative frequencies. In the other case, the word is just confusing the issue and the word ' probability ' there is not a good word at all.
>
> *Darwin*: You don't believe in betting.
>
> *Vigier*: I don't think betting is probability theory—it is amusing in itself.

One-concept theories, on the other hand, tend to ignore some of the important differences between the ordinary use of the concept and its use in science, and this I hope to show. It has traditionally been the task of philosophy to link up science with common sense, and to provide some sort of justification for the employment of scientific concepts by revealing their roots in everyday concepts. I wish to show in this

---

* Received 20.iv.65

paper that there is more unity in our talk about probability than is usually allowed by two-concept theories and more diversity than is allowed by one-concept theories.

We should start by taking note of an obvious but harmless ambiguity in the ordinary, everyday use of the word ' probable ', obvious but strangely ignored in discussions of the subject. The ambiguity is, moreover, indicative of one of the connections between the ordinary, qualitative, uses on the one hand and the technical, quantitative, uses of the concept on the other. The word ' probable ' is a concealed comparative, since it may be defined as ' more probable than not '. But we may see that the use of the word in the definiens is different from its use in the definiendum, if we try the experiment of replacing the word ' probable ' in the definiens by the whole definiens. We then get the nonsense ' more more probable than not than not '. Since this is nonsense, the two occurrences of the word must differ in meaning. This kind of ambiguity is not peculiar to the word ' probable '; it may be found in many scale-words, e.g. ' long ' in ' This is long ' and ' This is one inch long '. Now just as talk about the length of things follows naturally from talk about things being so-and-so many units long, so talk about the probability of an event being greater than that of another follows naturally from talk about one event's being more probable than another. The further step of making the concept quantitative is then a natural one.

Next we should notice that the probability-statements of our ordinary language are Janus-faced. Let us suppose that I have observed certain trends in the weather and I make a cautious prediction; I may say, ' It will probably rain tomorrow ' or ' It will rain tomorrow, probably ' or ' Probably, it will rain tomorrow '. Here my primary interest is not probability but rain. I insert the word ' probably ' parenthetically to make my assertion ' guarded ', as Toulmin puts it (*The Uses of Argument*, ch. II). We use the word ' probably ' to act as the signpost of a guarded assertion, just as I use the words ' I believe ' parenthetically (see Urmson, ' Parenthetical Verbs ' in *Essays in Conceptual Analysis* ed. Flew), when I say ' It will, I believe, rain tomorrow'. If somebody says to me, ' What you are saying is false ', he is not denying that I believe that it will rain tomorrow, he is denying what I believe, namely that it will rain tomorrow. If I wanted to, I could emphasise the words ' I believe ' or a similar form of words, so as to make it manifest that it was my belief which was the subject of conversation, by saying, for example, ' It is my belief that . . . ' or ' I do really believe that . . .'. Thus we should distinguish two uses of expressions like ' I believe ', the parenthetical and perfunctory and the non-parenthetical and non-perfunctory. The same kind of distinction is needed in discussing probability-statements. The Toulminian account is on the whole true of parenthetical probability-statements like ' It will probably rain '. If it does not rain, I shall be held to have been wrong. However undogmatic my assertion may be, I am not thereby invulnerable to refutation; my dogmatism or the lack of it is irrelevant to the truth or falsity of my assertion. This applies whether I safeguard myself with a ' probably ' or with a more tentative ' perhaps '. However, the Toulminian account is not even true of all parenthetical probability-statements. Thus we sometimes say things like ' He probably will, but possibly he won't '. Although the modal words ' possibly ' and ' probably ' are used parenthetically, we do not interpret this statement as containing two guarded but contradictory assertions at the same time. The limbs of these statements are not only compatible, but actually complement one another. They describe a situation which can be described with greater clarity, that is, in a way which leaves ' less scope for conjecture ', by a numerical probability-statement, that there is a probability of, say, 49/50 that he will. The ' but ' in these double-limbed parenthetical uses shows the other side of the same picture; logically it adds nothing, rhetorically it is important. The Toulminian account breaks down for these double-limbed uses, because the modal words are emphasised. Again, emphasis is given to the word ' probable ' when we say, ' It is as probable as not that it will rain ', ' It is more probable than not that it will rain ' and ' It is probable that it will rain '. If I make a non-parenthetical probability-statement like ' It is probable that it will rain ' and somebody retorts, ' What you say is false ', then he will be interpreted, not as denying that it will rain, but rather as denying that it is probable that it will rain. If, then, the Toulminian account is unsatisfactory, we have to investigate what counts against the truth of non-parenthetical probability-statements and so what is the relationship between probability, on the one hand, and truth and falsity on the other.

It is instructive in the first place to examine possibility-statements before considering probability-statements. The proper occasion for uttering the sentence ' It is possible that Higgins is dead ' is when one does not know but has some reason to believe that Higgins is dead. Now either Higgins turns out to be dead or he does not. If he turns out to be dead, the occasion for saying ' It is possible that he is dead ' has passed. Yet anybody who said without having knowledge, ' It is possible that Higgins is dead ' would have been making a true statement. In this case too *ab esse ad posse valet consequentia.* And since my previous felicitous utterance of ' It is possible that Higgins is dead ' has turned out to be true, I may now say felicitously and truly, ' It *was* (indeed) possible that Higgins was dead. I was right.' But let us suppose that Higgins turns out not to be dead. In this event the occasion for saying ' It is possible that Higgins is dead ' has again passed, at any rate for me, once that I know that he is not dead. However, I may discover that Higgins's house had been set on fire and that his life had been in danger and so on. This information would, I suggest, be sufficient to entitle me to claim justification and say, ' I was right. It was *possible* that Higgins was dead, but fortunately he escaped.' Let us now turn to probability-statements and see whether they are analogous. The proper occasion for uttering the sentence, ' It is probable that Higgins is dead ', is when one does not know, but has good though not conclusive reasons for believing that Higgins is dead. Now if Higgins turns out to be dead, the occasion for saying, ' It is probable that Higgins is dead', has passed. Moreover, one cannot infer here that if Higgins is dead, anybody who maintained that it was probable that he was dead, must have been right. There is no *ab esse ad probabile* principle to help us over this hurdle. But we need not despair. Whether we can justify our probability-statement *ex post facto* depends upon the circumstances. Let us suppose that Higgins's house received a direct hit from a conventional one-ton bomb. Nobody could have been reasonably expected to survive that. It was overwhelmingly probable that he was dead, and anybody who in such circumstances said, ' It is probable that he is dead ', would be judged to have made a true statement, indeed perhaps an understatement. If, on the other hand, Higgins turns out not to be dead, the occasion for saying, ' It is probable that Higgins is dead ' has once again passed. Let us again suppose that Higgins's house has received a direct hit, but that this time he is pulled out of the wreckage injured but alive. In this event we should say that he was extraordinarily lucky to escape, since there was ' every likelihood ' of his being killed. It was probable that he was dead and the chances of his survival were slender. If the above account of the use of expressions like ' It was probable ' is correct, then two theses of Toulmin's are mistaken. First, such expressions are not *Oratio Obliqua* in disguise. If we suppose that they are, then a wedge is driven between ordinary and numerical probability-language.

If ' probable ' means ' reasonable to believe ', as Toulmin for one seems to hold, then we surely sometimes want to say that it *was* reasonable for people to have certain beliefs rather than that it *seemed* reasonable for them to have them. Toulmin's second thesis, which he derives from the first, is that such expressions as ' improbable but true ' and ' probable but false ' are impermissible if not contradictory (op. cit. pp. 54-55). What, according to Toulmin, these phrases really mean is ' seemed improbable but true ' and ' seemed probable but false '. But if the account given here is correct, Toulmin's ruling is gratuitous. If I find that a statement turns out to be false, it does not follow from this that my statement that it was probable that it was true is false. For by saying that it was only probable that it was true, I implied that it was possible that it was false, and my having later found that it was false justified my original caution in allowing for the possibility of its being false. Toulmin's attempt to maintain that probability-statements are verified or falsified by the occurrence or non-occurrence respectively of the event predicted is due to his excessive concentration on single-limbed, parenthetical uses as opposed to the more complex but none the less common double-limbed and non-parenthetical uses.

I have maintained that the expressions which Toulmin alleges to be impermissible or contradictory are in fact neither and that the ordinary use is other than he says it is. But it does not follow from this that we ought to follow the ordinary use rather than Toulmin's. In fact, I suggest, it *would* be correct to interpret ' was probable ' as ' seemed probable ', if the universe were through-and-through Laplacean. If the universe were through-and-through Laplacean, it would be impossible, given the initial state of the universe or of any part of it, for any later state to be other than it is. Hence, after an event had occurred, it would always be false to say that it might not have occurred or that it had been possible that it would not occur. And if it were always false to say afterwards of any event that it was *possible* that it would not occur, it would *a fortiori* be false to say of it that it was *probable* that it would not occur. And if it were always false to say of an event afterwards that it was probable that it would not occur, one would always have to describe judgments of probability as having ' seemed true ' although false. Thus an observer in a Laplacean world, who knows that his world is Laplacean, ought to use probability-statements solely as expressions of his own ignorance and ought not to use ' it was probable ' with a meaning distinct from ' it seemed probable'. Hence, if anybody says, as Toulmin does, that 'it was probable' has no legitimate meaning distinct from the meaning of 'it seemed probable', he is prejudging a factual issue, viz., whether the universe is Laplacean, in the course of an apparently innocent conceptual enquiry. But in order to decide whether 'it was probable' can be used with a distinct meaning in our world, it is not necessary (though it is sufficient) that the world should be shown to be non-Laplacean. Our use of language is governed not so much by the way things are as by the way we think they are. Most of us regard a horserace as if we were in a non-Laplacean universe. Take Wisdom's example drawn from the Times Racing Correspondent (see 'A Note on Probability' in *Philosophical Analysis*, ed. Max Black): 'It was probable all Lombard Street to a China penny that Thankerton would win and then suddenly he stopped.' A Laplacean would argue that since Thankerton did stop, there must by the Principle of Sufficient Reason have been some knowable antecedent or characteristic of Thankerton which was responsible for his stopping, and if we had only known it, we should not have put our shirts on Thankerton. But few of us take this view of horse-racing. We do not think that the course of the race is predetermined. Not even God can help the bookies. When we say 'It was probable that Thankerton would win', we are saying not that people thought that Thankerton would win or that it seemed that any reasonable man would bet on Thankerton, but that any reasonable man would have betted on Thankerton. Whether or not our universe is Laplacean is a question which we need not decide here. The possibility of a non-Laplacean universe is sufficient to show that 'it was probable' has a meaning distinct from that of 'it seemed probable'.

What, then, makes us reject a non-parenthetical probability-statement? We reject a probability-statement, at any rate provisionally, when the predicted event fails on a considerable number of occasions. If on a large number of days when similar conditions prevail I say that it will probably rain, and on rather more than half of these occasions it fails to rain, then I have found reason to make at any rate a provisional rejection. This is an ordinary, non-scientific analogue of the more sophisticated procedures of rejection considered by Braithwaite in *Scientific Explanation.*

One might suppose from this that the ordinary concept can be adequately explicated in terms of frequency. Here again it is instructive to look at possibility. If I say that an event is possible, I can confirm my statement by showing that on at least one occasion an event of the same kind has occurred in similar circumstances. *Ab esse ad posse valet consequentia.* To refute the statement would be very much more difficult, since we should have to show that the event never could happen. From the above considerations one might be led to suppose that ' probable ' : ' possible ' :: ' usually ' : ' sometimes '. If this were so, it would seem that a frequency theory of probability would be tenable if and only if a frequency theory of possibility was tenable. But a frequency theory of possibility is implausible. Though a statement of possibility is conclusively verified by finding that at least one event of the kind has occurred, this is not what a possibility-statement means. This comes out clearly in the case of logical possibility. For to establish logical possibility an empirical investigation is not necessary, although it is, of course, sufficient. A logical-possibility-statement may be true and yet the possibility in question may never be actualised. A logical-possibility-statement, then, cannot be identical with a statement about frequencies. Now if there is a concept of probability in the same family of concepts as logical possibility, this also cannot be explicated in terms of frequencies. That there is such a concept, in a position formally analogous to that of the ordinary concept, is shown by its employment by logicians and epistemologists, although it is not used in our everyday language. When philosophers discuss the nature of human knowledge, they often want to say things like, ' All our knowledge is only probable ', meaning by this that there is a logical possibility of our being mistaken. Since this use of ' probable ' is complementary to the use of ' logically possible ', we have here an instance of the logical concept.

Nor is a statement of non-logical or empirical probability identical with a statement about frequencies. For a statement of empirical probability may be true despite the fact that not one event of a certain kind has occurred—this is ' probable ' 's analogue of the principle *a posse ad esse non valet consequentia.* It is always logically possible for the probable not to happen. Nevertheless there is a special relationship between empirical probability and frequency, a special relationship which accounts for the continuing search for a satisfactory frequency theory of probability. Indeed, if we restrict our attention to finite classes, the relative frequency may be identified with the probability, and so one might suppose that one could explain the *whole* meaning of the word ' probable ' in terms of frequency. But since the word ' probable ' is an evaluative word, this would be to commit an analogue of the so-called ' Naturalistic Fallacy '. But, *pace* Toulmin, it is no ordinary contingent fact that we think that the usual is to-be-expected and the rare is not-to-be-expected, that it is rational to expect the usual and irrational to expect the unusual. What is rational or irrational depends on one's aims and purposes. It is our aim in making statements and predictions about the world to ensure that as much as possible of what we say should be true. Now if we expect the unusual, we shall be more often wrong than right and thus fail to further our purpose. If you are confronted with a bag which you know to contain 100 balls, 95 of them white and 5 black, it would be irrational to expect to pick out a black ball on your first draw. Since we want to be more often right than wrong, it is rational to expect the usual and irrational to expect the unusual. Because of this we are justified in using the word ' probable ' as a term of commendation. In this way we can both have our Naturalistic cake and eat it.

I want now to examine what seems to be an important distinction between numerical and non-numerical probability-statements. On the one hand, we say things like, ' The probability of an A being B is $\frac{3}{4}$ ', while on the other we say, ' It is probable that X is a B ' (where X is known to be an A). One should not, I think, gloss over this difference, as some philosophers do when they say that the second way of talking is an elliptical version of the first way of talking. Thus Keynes regards probability as ' relative to evidence ' and all statements of the form ' It is probable that *p* ' as elliptical. Since probability is supposed to be a logical relation, statements of the form ' It is probable that *p* ' cannot be refuted merely by looking at the facts, but by examining the relation between the unmentioned evidence and the conclusion *p*. Now Keynes and his followers were quite right in thinking that evidence does make statements probable and that there is a relation of making or rendering probable, of ' probabilification ' (Kneale). But it does not follow from the existence of such a relation that any statement of the form ' It is probable that *p* ' is an elliptical way of saying ' p is probabilified to a degree over $\frac{1}{2}$ by evidence '. The second statement may well be our reason for asserting the first, but it does not follow from this that the first is just shorthand for the second. Compare the case of entailment. In this case ' p ' does not mean the same as ' q entails p ', nor when we say ' It is necessary that *p* ' are we saying elliptically ' q necessitates p '. Probability is compared by Keynes with distance. ' No proposition ', he says (*Treatise on Probability*, p. 7), ' is itself either probable or improbable just as no place can be intrinsically distant '. When we say that a place is three miles distant we mean that it is three miles ' from where we are then situated or from some starting-point to which we tacitly refer '. But there are surely some important differences between probability and distance. When I say, ' X is 3 miles distant ', it is usually plain that it is three miles distant from the place from which I am speaking. In the case of probability, the evidence I have at my disposal is not manifest in the way that the place of utterance usually is. When you say to me ' It is probable that *p* ', what you say does not mean that *p* is probable on the evidence at your disposal, evidence to which you ' tacitly ' refer. For you may not have a shred of evidence for your statement, it may be a guess or a hunch. Again, in the case of distance I do not contradict you if I say that Perth is 200 miles distant, and you say that Perth is 20 miles distant, for we may be speaking on the long-distance telephone. But if I say, ' It is probable that it will rain in Dundee tomorrow ' and you say ' It is probable that it won't rain in Dundee tomorrow ', (i.e. ' It is improbable that it will rain ') we are contradicting each other, whether we have the same evidence or not. But if the Keynesian analysis were correct, we should not be contradicting one another, but making true, though elliptical, statements about the logical relations between our respective evidence-statements and the statement ' It will rain tomorrow '. Further, one of us might be persuaded to withdraw his probability-statement, when presented with further evidence. But if the sentence ' It is probable that it will rain ', as uttered by the meteorologist, were really elliptical for a probabilification-statement, he would not need to withdraw it, he would only need to put it away in the archives. There is thus an important difference between probability-relation statements and statements of the form ' It is probable that *p* '. (See further A. J. Ayer, *The Concept of a Person and Other Essays*, No. 7). Probability-statements may, I suggest, be regarded as the conclusions of probability-arguments having the form ' *p*; so it is probable that *q* '. Probabilification-statements would then function as principles of probability-arguments, clearly distinguished from the conclusions of such arguments. (Toulmin makes a similar point in his book, op. cit. p. 81). We could take into account those features of probability-statements which made Keynes call them ' elliptical ' and ' relative to evidence ' by saying that they are ' detached '. I call a statement ' detached ' if it is implied by another statement which is asserted, and so enables us to assert the implied statement in its own right. It is not the case that all our probability-statements are detached; some of them are not based on any evidence and this is a defect. It is a requirement of rationality in our probability-talk that all our probability-statements should be detached, that we should not insert the word 'probably' gratuitously into our hunches and guesses. Now non-detached ('baseless') probability-statements have no 'descriptive meaning' and may be said to be genuinely subjective. Are we to say, then, that there is a distinctive subjective concept of probability? We could say that there was, but this would obscure the fact that there is a continuum of uses of the word 'probable' ranging from the purely objective to the purely subjective. The word 'probability' is used purely objectively when it is used as a synonym of 'relative frequency' or 'the limit of the relative frequency'. A purely subjective use has only evaluative meaning. By saying that an event is probable (subjectively), all that we imply is that the event is to-be-expected, we are expressing a groundless belief that the event will occur. Such subjective uses are best expressed parenthetically. It would seem, then, that what both subjective and objective uses have in common is membership of a continuum of uses, the central members of which have both evaluative and descriptive meaning. Once this is seen, there is no need to labour the question whether we have here one or two or more concepts of probability. It is clear, however, that the diverse theories of probability (e.g. Personalist and Frequency) take their rise from emphasis on opposite ends of the continuum.

However, it might be objected, there is a limit to this straining after unity. Surely, it might be said, one big difference remains between numerical and ordinary probability-statements. In making numerical probability-statements we are concerned with ratios between classes or relations between propositional functions. But the probability statements of our ordinary language are frequently concerned with particular cases. As Russell puts it (*Human Knowledge*, p. 358): 'When I say that the chance of a coin coming heads is a half, that is a relation between two propositional functions "x is a toss of the coin" and "x is a toss of the coin which comes heads". If I am to infer that in a particular case, the chance of heads is a half, I must state that I am considering the particular case solely as an *instance*. If I could consider it in all its particularity, I should in theory be able to decide whether it will fall heads or fall tails and I should no longer be in the domain of probability.' When you say to me, 'You will probably live to be seventy', you seem to be talking about this particular case and not merely talking about me as a member of a particular class. But when we give a numerical value to the probability, things are different. We may assert any proposition of the form 'The probability of a B being an A is m/n', but we cannot say *without qualification* 'The probability of *this* B being an A is m/n'. Now the probability of my living to be seventy will vary according to the class which I select and I have a large choice, since there are an indefinitely large number of classes of which I am a member. The larger the number of relevant characteristics I include in my description of the class, the more satisfactory as a guide to action will be the numerical probability which I reach. 'Every circumstance of my health and my way of life is relevant, but some of them may be so uncommon that I can get no reliable help from statistics' (Russell, op. cit. p. 358). The reason for this uncommonness is of course, that human beings are not mass-produced automata, programmed in accordance with a rigid rubric, but individuals with widely different characteristics. If I included all relevant characteristics of myself in the description of the class, the only member of the class would be myself and thus I could not obtain any statistics. At first blush, then, it looks as if I can give no sense to unqualified singular probability-statements like '*The* probability of my living to be seventy is ½', in which no mention is made of the class of which I am to be considered as a member.

But a meaning can be given to *unqualified* probability-statements in the following way. Let us consider a non-Laplacean universe, containing a class of objects indiscriminable from each other in all respects except spatially. Of these, let us say, 1 per cent every year disappear, but since there are no discriminable differences between the individual objects, there are no factors which make it more likely that any one should disappear rather than any other. In this case, the judgment that the probability that any one object should disappear in the next year is 1/100 would be the final judgment even of omniscience. With the fullest possible information about the object 'in all its particularity' one cannot decide whether the object will disappear or remain, one cannot make an exit from 'the domain of probability'. The logical possibility of this imaginary situation is all we need to give a meaning to unqualified probability-statements. In fact we often behave as if the world were non-Laplacean and this is quite sufficient to warrant the use of unqualified probability-statements. For, as remarked before, our use of language is determined not merely by the way things are, but also by the way we think they are.

Further, a meaning can be given to *singular* probability-statements in the following way. To consider me 'in all my particularity' is to consider me as a member of a unit class with the consequence that no statistics would be obtainable. There is one way, however, in which we can circumvent this obstacle, and that is by considering what *would* be true of a class of individuals exactly like the one under discussion in all relevant respects. I suggest, then, that a singular probability-statement which is allegedly non-statistical is really founded on a conjecture as to what proportion of individuals, having exactly the same characteristics as the one under discussion, would have some other property, say, that of surviving until the age of seventy. So when the doctor examines me and says 'You will probably live to be seventy', what he means is something like, 'If there were a class of individuals exactly like you in all relevant respects, then the relative frequency of those living to be seventy would tend increasingly to exceed $\frac{1}{2}$ the greater the number of members of the class'. If this is what is meant, we can understand and account for the disrespect or rather contempt shown by many for 'unscientific', 'qualititative', 'subjective' probability-statements. (For examples see P. Meehl, *Clinical vs. Statistical Prediction.*) Counterfactual conditionals are notoriously difficult to verify or support and this is why singular probability-statements are so frowned upon. Further, on this account we can see how right Peirce, von Mises and Popper were to regard the probability of an object's undergoing a change as a property of the object, as a 'would-be' or a propensity. As long as an object is the member of a unit class, its 'would-be' lies hidden; it can only be made manifest in an ensemble. Sir Charles Darwin was 'absolutely certain that' the horse was not going to run 100 times. But he was not justified in inferring from this that we must have an 'independent, separate idea of probability'. Nor does it follow from this that in order to talk of the probability of the horse's winning we must be able to calculate relative frequencies, as Vigier maintained. It is only a *contingent* fact that the horse is not going to run 100 times and this fact, though a bar to our 'calculating relative frequencies', is no bar to our trying to conjecture or imagine how the horse would fare if it were to run in 100 races. This may be an 'unscientific' use of probability, but it is nonetheless an intelligible one related to the 'scientific' one. We are not compelled to throw up the sponge and declare either that we 'have got to have an independent, separate idea of probability' or that in the case of betting 'the word "probability" . . . is not a good word at all'. As I have remarked, singular probability-statements have sometimes been pooh-poohed as 'subjective', and I agree that they have not the scientific 'respectability' of statistical probability-statements. I should prefer, however, to call them 'impressionistic' and to reserve the word 'subjective' for 'non-detached', 'baseless', probability-statements.

Lastly, I want to show that there is no need to postulate a special inductive kind of probability, although some account has to be given of the use of probability-language in inductive contexts. When we have a certain numerical proportion m/n of B's which are A's in a finite population, we can infer without any mediating principle that the probability in this population of a B being an A is m/n. Now when we are making an inference about the future, we are 'given' a past population and past samples, but the future population is not 'given'. All that we can deduce is that *if* the future population matches the past population, then the probability of a B being an A is m/n. We cannot deduce unconditionally what the probability is; we are dependent in our inferences about the unknown on some leading principle like the inductive hypothesis, we are dependent on the assumption that future populations will match past ones. It is because of this inductive leap that we have the illusion that we must be dealing with a distinct kind of probability, inductive probability.

We have, then, reached the following conclusions:

(1) The everyday and numerical uses of the concept of probability are linked.

(2) There is a logical concept of probability whose use is limited to philosophical contexts and cannot be explicated in terms of relative frequency.

(3) The empirical, everyday concept of probability is logically connected with the concept of relative frequency, but its meaning is not exhausted by the latter concept.

(4) The rationality of commending the usual as 'probable' has been vindicated.

(5) Subjective and objective uses of probability-language have been distinguished and it is suggested that they form part of a continuum of uses which gives the empirical concept of probability its unity.

(6) A meaning has been given to probability-statements which are both unqualified and singular.

(7) The alleged existence of a concept of inductive probability distinct from the empirical concept is rejected.

Queen's College, Dundee,
University of St Andrews

DISCUSSIONS

In Defence of the Sociology of Knowledge

Professor I. C. Jarvie has chosen to review my *Philosophy, Science and the Sociology of Knowledge* in what must be considered an odd way for a scientific journal.[1] He chooses to argue with the book, next argue with the sociology of knowledge in general, and finally dismisses the whole field as if it were phrenology or alchemy—something 'dead' and 'better left to rest in peace'. This seems to be unseemly haste to do away with what is only now becoming a central issue in the philosophy of science : the place of ideology in the formation of exact knowledge. Mr Jarvie's rhetorical question : 'Biases or total ideologies, could be socially or economically determined, I suppose ; but how can *knowledge* be?' makes two dubious assumptions. First, it utterly fudges what the sociology of knowledge is about—the interpenetration of ideological factors in real knowledge. Second, the reviewer seems much clearer than any philosophy of science scholar I have ever spoken to, known of, or read, as to what are exactly the constituents of 'knowledge' (in italics no less).

The reviewer speaks of theories as equivalent to opinions! And actually implies that one can develop a sociology of knowledge only if by knowledge we mean theories. Surely he does not seriously mean that knowledge is capable of being organised in a non-theoretical fashion? Or that knowledge is identical to fact, or that knowledge is logical relations. For if the former, we do not need his review, he could simply point to the book with a thumbs down sign. While if the latter, he is simply engaged in an idiosyncratic mood, whereby logical truth is equated with empirical information about the world.

Mr Jarvie is honest enough to admit that he has a strong disposition against the sociology of knowledge—which the book apparently did not dispel. But he is very unfair to make it appear as if this was the task of the book to begin with. I did not write a manual for non-believers (although this clearly could be done with some profit) ; but rather a guide through the literature : that is why the book has the compressed, abstract and low-on-examples qualities the reviewer complains about. The volume further has an ecumenical intent : of bringing philosophers of science and sociologists of philosophy closer together ; and at the epistemological level, of working to remove the obstacles between positivist and naturalist styles of doing social philosophy.

Mr Jarvie 'deplores' the sociology of knowledge because it fails to take seriously the question : how is the sociologist of knowledge specially privileged in not being biased by *his* total ideology? After acknowledging that I 'affect to deal with this criticism', Mr Jarvie fails to say how I dealt with it. True enough, I contrast my position with that of Mannheim. First, because Mannheim creates a Platonic sociologist who stands above the cares of the world in order to see the truth of the world ; and second, because Mannheim's relativism or relationism leaves the problem of the sources of thought quite intact. My own answer, as given in the book, goes as follows : (*a*) The sociologist of knowledge is first of all a sociologist, and hence shares in a general methodological orientation with other sociologists; (*b*) the sociologist of knowledge explains the relation of ideas, opinions, thoughts, theories, etc., to their material sources, but in turn leaves for the sociologist of sociology to explain his role, which admittedly is close to relativism, to the infinite elasticity of thought ; (*c*) the sociologist of knowledge makes no assumption of the 'total' character of an ideology. Indeed, some professional ideologies may be as significant as Statist ideologies in given circumstances. Also, ideology may be fragmentary, eclectic, national, etc.; (*d*) the sociologist of knowledge makes no assumption about the good or evil of an ideology ; indeed, he takes seriously the ways in which ideology is a stimulus to science no less than theology ; (*e*) the sociologist of knowledge leaves open the possibility that scientific knowledge is relatively autonomous, especially after a certain intellectual 'take-off' period ; (*f*) the sociologist has a general disposition to connect such scientific knowledge, or religious beliefs, etc., with other variables. In this way, the study of knowledge is much like the study of bureaucracy, stratification, race relations : an examination of how the critical variable is plugged into or fused with other social factors. This is not to minimise the special problems arising out of the study of information or belief systems, but it at least serves as a clear warrant and legitimation for such kind of work—which sociologists choose to call the sociology of knowledge.

---

[1] See this *Journal*, 1964, 15, 266–268

The book is now five years old. There are many points at which I find much to fault the book on. The book is indeed too abstract. It needed more study of the sociology of scientific theory as the reviewer suggests. And in a series of papers during the past half decade, myself, along with a number of colleagues, and now some students as well, are beginning to answer the pressing questions of how human beings come by their ideas : how they choose their career lines, the points of view they have, the effect of immediate university environment on decisions, even studies of personal and family backgrounds as key variables in career choice and ideology base.[1] But of course, to all of this, Mr Jarvie will undoubtedly say, as he did in his review, that this 'is of no epistemological interest'. To which I would reply : this attitude is what has made most philosophical cookies crumble.

I do not much mind being convicted of 'sharing a Bacon-Locke view' (horrors!). I can also appreciate the reviewer's gambit of telling me what I should have written instead of what I did write. I can even appreciate his 'reluctance' to accept my attempt to 'rehabilitate' the sociology of knowledge. But I can much less appreciate the accusation of dealing with big problems and big theorists in a large-scale way. The reviewer accuses me of employing '(*a*) through (*g*)' definitions of the sociology of knowledge. It is kind of him to provide the list for the less logically precise reader.

---

[1] The first fruits of this on-going co-operative effort can be found in a volume which I edited recently entitled *The New Sociology*, London and New York, 1964. See in particular the papers by Anatol Rapoport, Ernest Becker, Abraham Edel, Alvin W. Gouldner, and Sidney M. Willhelm. In addition, I have written various papers intended to deepen the study of the sociology of knowledge—along several lines—the connection of ideology to political systems, and the value orientations of social scientists. See, 'A Formalization of the Sociology of Knowledge', *Behavioral Science*, 1964, **9**, 45-55; and 'Professionalism and Disciplinarianism: Two Styles of Sociological Performance', *Philosophy of Science*, 1964, **31**, 275-281.

Yet, it would have been nicer still had Mr Jarvie noted that what he lists are not definitions, but instead, *descriptions* of, *theories* about, and *tasks* for, the sociology of knowledge. Further, although there is the clever inference that I have employed multiple definitions of the same object, my critic fails to note that what is being examined is not an object (like a brick or even a colour), but a general sub-branch of a major social science, sociology—which unfortunately, Mr Jarvie shows not the slightest inclination to take seriously.

Irving Louis Horowitz

Washington University (St. Louis)

Reply to Professor Horowitz

I am sorry Professor Horowitz did not like my review of his book,[1] ad it is understandable that he prefers to write his own, with its up-to-date summary of other work-in-progress. For one thing, after summarising the book, I chose ' to argue with ' it and Professor Horowitz considers this ' odd ' in a scientific journal. For my part I consider extended controversy over reviews out of place in *any* journal, so I shall be brief. Two critical points seem worth the further attention of readers of this *Journal*: (1) Can there be ' interpenetration of ideological factors in real knowledge '; and (2) Can the sociologist of knowledge aspire to ' relativism, to the infinite elasticity of thought '?[2] While not prepared to agree that ' the place of ideology in the formation of exact knowledge ' is ' becoming a central issue in the philosophy of science ', I do think (1) and (2) of some interest to philosophers of science.

(1) Whatever knowledge is, and I am sorry Professor Horowitz feels I am too confident about it, it is not the same as the statements which try to capture it. On the other hand ideologies are statements (and false ones at that). How then can false statements interpenetrate knowledge? In my review I said that theories (statements of putative knowledge) are what should really be under discussion.

(2) Professor Horowitz's lengthy evasion of the problem raised by the ideology of the sociologist of knowledge will not do. The ideas, theories, analytical tools of the sociologist can be, according to the sociology of knowledge, interpenetrated with ideological content. Thus what the sociologist reports about what he is studying may be distorted. Has then his work no claim to truth? Everything he says may be distorted, and it does no good to shove the problem onto someone else trying to do the sociology of knowledge of sociologists of knowledge, as Professor Horowitz suggests in his point (*b*). This would lead both to an infinite regress (of students studying the students) and would fail to get rid of ideology, which increases at each level since both subject studied and student studying are 'interpenetrated'. The defence that ideologies may be only partial or eclectic (Horowitz's point (*c*)) is not effective, since the *possibility* of ideological distortion is all that needs to be established to throw doubt on the whole enterprise. (As soon as we can tell when bias is or is not present the problem of bias, and thus of the sociology of knowledge, disappears.)

---

[1] Although if he had read it more carefully he might have avoided such slips as accusing me of accusing him of employing multiple definitions which I nowhere do.

[2] Leaving aside the obvious point that any statement of relativism is paradoxical, in the strict sense.

The choice is clearly between the sociology of knowledge and the claim to truth, the one excludes the other.[1]

While denying his book was a manual for non-believers (in the sociology of knowledge), Professor Horowitz claims the ecumencial intent of 'bringing philosophers of science and sociologists of philosophy closer together'. As a non-believing philosopher of science I was eager to find out from this ' guide through the literature ' what I had been missing and should now take an interest in. I stick to the conclusion of the review that this desire was not gratified by *Philosophy, Science and the Sociology of Knowledge.*

I. C. JARVIE

NOTE ON MR KESWANI'S ARTICLES, ORIGIN AND CONCEPT OF RELATIVITY

Mr. G. H. Keswani's articles[2] on the origin of the ideas now loosely grouped and very much confused under the name, ' relativity ', are very valuable contributions to an important subject. They are not merely of historical interest; they have a vital bearing on problems of current concern, for I think there can be no doubt that much of the misunderstanding shown in recent discussions of this subject arises from a general unawareness of what actually took place during the first two decades of this century. The numerous references that Mr Keswani gives should be of great help to those who really wish to get at the truth.

I think, however, that he has failed to reach the heart of the matter by presenting it as a question of *priority of discovery* of some unique thing called ' the principle (or postulate—the distinction is insignificant) of relativity '. In fact the situation was quite different, and much more complex. ' The principle of relativity ' had various meanings, and the theories associated with it were quite distinct; they were not different forms of the same theory. Each of the three protagonists considered by Mr Keswani—Lorentz, Poincaré and Einstein (he does not deal with the fourth, Ritz, whose ' relativity ' was of a still different variety)—was very well aware of the work of the others (if not at the times of writing their papers, then very soon afterwards) but each preferred his own views, and it would not have occurred to him that any question of ' priority of discovery ' entered into the matter. So when, as for example, Mr Keswani points out, ' nowhere did he [Poincaré] credit Einstein with the discovery of the theory of relativity ', the significance of this is not that Poincaré erred in a matter of priority, but that ' the theory of relativity ' meant to him the theory of Lorentz, and not that of Einstein; and when, long after Poincaré's death, Lorentz in 1928 said, ' the theory of relativity is really solely Einstein's work ', he was speaking at a time when ' the theory of relativity ' had come uniquely to denote Einstein's theory, as the remarks immediately preceding this comment, which Mr Keswani quotes, clearly show. To take a single example of Poincaré's meaning—so late as 1912, in discussing Einstein's ideas concerning the action of light on molecules, he says that in this problem we must apply ' le principe de relativité de Lorentz '.[3] It is inconceivable that at that date, and in such a connection, he would have so written had he not considered Lorentz's theory to be essentially different from Einstein's.

---

[1] With respect to the narrower problem of *Wertfreiheit* this issue is cogently set out in K. Klappholz, ' Value Judgments and Economics ', this *Journal*, 15, 97 et sqq.

[2] ' Origin and Concept of Relativity ', this *Journal*, 1965, 15, 286–306, 16, 19–32

[3] *Dernières Pensées*, Chap. 7

The matter is still further complicated by the fact that 'the principle of relativity' not only meant different things to different people, but, for Poincaré, it changed its meaning as time went on. I think Mr Keswani is in error in saying, in a footnote, that at any time, 'Poincaré categorically denies the aether'. The reference he gives does not bear this out, and, in fact, Poincaré could not have done so without denying the whole of electromagnetic theory, which he certainly never did. The only way of retaining electro magnetic theory without an ether is that of Einstein, which, in fact, had not been published at the time (1903) when, in *Science and Hypothesis*, Poincaré wrote the passage which Mr Keswani construes as a denial of the ether. And that he did not abandon the ether later is shown by the title of the 1912 address from which I quoted in the last paragraph, viz. 'Les Rapports de la Matière et de l'Éther'.

In *Science and Hypothesis* (published before Lorentz's derivation of his complete transformation, but after the suggestion by FitzGerald and Lorentz that bodies moving through the ether were contracted), Poincaré means by the principle of relativity simply what is implied by Newton's first law, that the laws of motion are independent of the inertial co-ordinate system (to use a later term) to which they are referred. 'The movement of any system whatever', he says, 'ought to obey the same laws, whether it is referred to fixed axes or to the movable axes which are implied in uniform motion in a straight line' (Chap. 7). After Lorentz's 1904 paper, however, he gave it the new meaning that 'we shall never succeed, by any imaginable method, in demonstrating any but relative velocities',[1] and his discussion of the principle shows that he certainly believed that a body *had* a motion through the ether, but that the effects of that motion combined to make it undetectable by anyone confined to the body. His final word on the subject[2] clearly expresses his change of view: 'Quelle est alors la révolution qui est due aux récents progrès de la Physique? Le principe de relativité, sous sa forme ancienne, a dû être abandonné, il est remplacé par le principe de relativité de Lorentz.' Never at any time does he apply the term 'principle of relativity' to Einstein's 'postulate of relativity', viz. 'the phenomena of electrodynamics as well as of mechanics possess no properties corresponding to the idea of absolute rest'. To Poincaré they certainly did possess such properties, but we could never detect them.

There were thus, in the period before the first world wär, three distinct ideas denoted by the term, 'principle of relativity' (all of course, antecedent to the *special theory* and *general theory* of relativity); they are: (1) That the laws of motion recognise no distinction between rest and uniform motion (Newton, Ritz, and early Poincaré); (2) That the laws do recognise such a distinction, but the real (not merely conceptual) effects of motion are such as to make it impossible for any one to determine his own motion without reference to outside bodies (Lorentz and later Poincaré); (3) That there exists no natural standard of rest that would make it meaningful to say that any single body has one motion rather than another (Einstein). When this is clearly understood, it is seen that the question, 'Who first discovered the principle of relativity?' is meaningless. There is no one principle of relativity; none of the conceptions bearing that name has been, or can be, ' discovered '; and we still do not know which, if any, accords most satisfactorily with natural phenomena.

---

[1] *Science and Method* (1908). English translation by Francis Maitland (Nelson and Sons; no date); p. 221.

[2] *Dernières Pensées*, Chap. 2

The importance of all this for current discussions can hardly be exaggerated. Many writers seem unaware that there are even two different ' relativity theories ' (Lorentz's and Einstein's), and on different occasions they appeal to one or the other, according to the need of the moment, without consciousness of the inconsistency. For example, those who maintain that, on Einstein's theory, asymmetrical ageing is ' real ', refuse to accept my conclusion[1] that that theory is untenable, on the ground that the ageing is only an appearance depending on one's choice of co-ordinate system.[2] One of the most naive, but nevertheless quite common, confusions is the idea that if you call an absolute standard of rest, ' the ether ', you have Lorentz's theory, but if you call it ' the universe ',[3] you have Einstein's. Since Einstein's postulate says that there is nothing in electrodynamical or mechanical phenomena corresponding to absolute rest, this has the necessary implication that some other properties of the universe qualify it to act as an absolute standard of rest, but I have never seen these specified: can they be psychic? Others, who recognise that the theories are distinct, yet regard them as experimentally indistinguishable; Mr Keswani cites von Laue (1911) and Margenau and Mould (1957) to this effect. If that were so they would simply be different verbal expressions of the same theory, which they certainly are not. Others again accept Einstein's special postulate of relativity, but regard its generalisation to all motion as illegitimate.

A proper understanding of the history of the subject would at once clear away all this confusion. To take the last point first, it is obvious that Lorentz's principle of relativity cannot be generalised, whereas Einstein's not only can but must be. There is no passage at all from the postulated effects of uniform motion through the ether, which combine to disguise that motion, to effects that will disguise all types of accelerated motion. Some independent new theory would be required for that, and no reason can be imagined why acceleration should not be absolute; as Mr Keswani points out, Poincaré regarded it as self-evident that it was. On the other hand, Einstein's relativity—that there is no distinction (not merely no *detectable* distinction) between any two states of uniform motion of a single body—necessarily implies that a change from one such state to another can have no meaning, for the body can equally well be credited with either state both before and after the ' change '. Accordingly, Einstein *had* to generalise his postulate: he saw this clearly and, as we know, he did generalise it. Hence, anyone who holds the ' special ' but not the ' general ' postulate of relativity necessarily holds to Lorentz's theory and not Einstein's, whether or not he realises that fact.

As for experimental distinction between the theories, that is easy enough to find when one understands in what the distinction consists. It is not in the mathematics but in the physical ideas which the mathematics represents.[4] Any experiment that merely tests the equations, irrespective of the meanings assigned to the symbols in them, will of course fail to provide a distinction, for if it verifies (or falsifies) Lorentz's theory with his meanings for the symbols, it will automatically verify (or falsify) Einstein's theory with *his* meaning for the symbols. But there is no lack of possibilities of a *physical* distinction; the simplest (in principle) is to test for asymmetrical ageing. This is required by Lorentz's theory, in which the slowing down with motion is, as he put it, as real as the contraction of a body through cooling. On Einstein's theory, however, on which the motion can be assigned indiscriminately to the Earth or to the travelling clock, it cannot affect one differently from the other.

---

[1] *Nature*, 1962, **195**, 985–6; this *Journal*, 1964, **15**, 41–61

[2] E.g. M. Born, *Nature*, 1963, **197**, 1287

[3] E.g. F. S. Crawford, Jun., *Nature*, 1957, **179**, 1071-1072

[4] Whittaker (*History of the Theories of Aether and Electricity*, Vol. II, 1953) unfortunately failed to distinguish between the two theories, and when his researches, aided by his memory, assured him that Lorentz and Poincaré had indeed constructed a theory that mathematically was identical with that then known as ' the theory of relativity ',

For a reason which I do not understand, this last point seems to be beyond the comprehension of most writers on this subject. In an attempt, which I thought could hardly fail, either to make it clear or else make clear some basic misconception of my own, I put it some years ago into the form of a single syllogism,[1] so that either the conclusion could be generally accepted, or my error located in one of the three items. To my astonishment, however, neither issue was attainable. All my efforts, then and subsequently, have failed to produce one single relevant comment on the syllogism; indeed, without the absence of such comment none of the numerous papers on the subject seems to be complete. The syllogism has borne the pageant of its bleeding heart through journal after journal,[2] but, alas, there is none who will respond to its pleadings. Yet just that one simple thing would serve to settle at a stroke a controversy that has raged for more than fifty years. Is it possible to hope that peradventure there may yet appear one who is sufficiently aware that open loyalty to truth, even though less attractive, is more honourable than silent commitment to orthodoxy, to do just that little thing? If so, that will be the day!

Just a minor point in conclusion. In his first paper, Mr Keswani states that the velocity of light compounded with a velocity v remains the same even if v is not in the

---

he credited them with the authorship of it. This was a much needed corrective to the view, then current, that 'the theory of relativity' was Einstein's unquestionable possession, but it failed to give Einstein credit for his own original work, which was different from that of Lorentz and Poincaré. As a mathematician, Whittaker would be particularly susceptible to the danger of judging a theory by its mathematical form alone, but it is most unfortunate that on this occasion he did not sufficiently guard against it.

[1] *Nature*, 1957, **179**, 1242

[2] Here are some of its repetitions: *Discovery*, 1957, **18**, 174; *Science*, 1958, **127**, 158; *New Scientist*, 1961, August 24; *A Threefold Cord*, by Samuel and Dingle (Allen & Unwin, 1961). Shamefacedly I repeat it once more here, since, if there should be anyone who can brace himself to face it, he eminently deserves to be saved the trouble of looking it up elsewhere:

(1) According to the postulate of relativity [Einstein's, of course] if two bodies (for example, two identical clocks) separate and reunite, there is no observable phenomenon that will show in an absolute sense that one rather than the other has moved.

(2) If on reunion one clock were retarded *by a quantity depending on their relative motion*, and the other not, that phenomenon would show that the first had moved and not the second.

(3) Hence, if the postulate of relativity is true, the clocks must be retarded equally or not at all: in either case, their readings will agree on reunion if they agreed at separation.

same direction as the light. But any theory that required that would be unable to explain aberration: it is not required by the Lorentz transformation.

HERBERT DINGLE

THE ORIGIN AND CONCEPT OF RELATIVITY

IN previous issues of this *Journal*, [1,2] G. H. Keswani has attempted to prove that Einstein's Special Theory of Relativity is inconsistent. Professor Born pointed out [3] that Mr Keswani had made an elementary mistake in his calculations. However, in his reply Keswani quoted two instances from Einstein's 1905 paper which seem to show that Einstein himself is inconsistent. Does this, therefore, mean that Special Relativity is to be abandoned? My contention is that both Einstein and Keswani have made the mistake pointed out by Born; but since they start from slightly different premises, Special Relativity is saved.

Essentially the controversy is as follows. Keswani assumes the Principle of Relativity and the constancy of the velocity of light for all frames moving uniformly with respect to one another. He considers two such frames S and S' and concludes that a spherical wave-front observed in S, will be seen as a spherical wave-front by an observer in S'.

From this he derives the Lorentz transformation and applies it to the wave-front in S. He concludes that the observer in S' regards the wave-front as an ellipsoidal one, which is contradictory to his first conclusion.

Professor Born points out the error.[3] The time of an event in S' depends on both the *position* and *time* of the same event as seen in S. Hence the observer in S' does not regard all the points on the transformed wave-front as occurring at the same time. It follows that a *particular* wave-front in S is not regarded as a wave-front in S'.

In his reply, [2] Keswani quotes Einstein that,

(i) . . . the wave under consideration is therefore no less a spherical wave with velocity of propagation c when viewed in the moving system . . .[4], and that

(ii) . . . the spherical surface—viewed in the moving system—is an ellipsoidal surface. . . .[5] It appears that Einstein himself is inconsistent.

In his analysis Einstein considered a spherical wave-front;

$$x^2+y^2+z^2-c^2t^2 = 0. \quad (1) \text{ in a frame } S.$$

On applying the Lorentz transformation to a frame S' moving uniformly with respect to S, he obtains the result:

$$x'^2+y'^2+z'^2-c^2t'^2 = 0 \tag{2}$$

where (x, y, z, t) and (x', y', z', t') are the co-ordinates of an event as seen in S and S' repectively. *Apparently* the transformed surface is also spherical. This is not the case. Although t is constant over the surface (1), t' is not constant over (2), since t'

---

[1] *The British Journal for the Philosophy of Science*, 1960–61, **11**, 50

[2] Ibid. 1961–62, **12**, 151

[3] Ibid. 1961–62, **12**, 150

[4] *The Principle of Relativity* (Dover), A. Einstein. 'On the Electrodynamics of Moving Bodies', p. 46

[5] Ibid. p. 57

depends on both the *position* and *time* of events on (1). Einstein has made the same mistake as Keswani.

Einstein reaches conclusion (ii) by attempting to relate the energy per unit volume of a light beam in one system, to its observed value in a system moving uniformly with respect to the first.[1] This he does by considering a spherical volume of space moving along the beam at the speed of light, that is the volume enclosed by the surface;

$$(x-lct)^2+(y-mct)^2+(z-nct)^2 = r^2 \qquad (3)$$

where (l, m, n) are the direction cosiness of the light beam. This volume must, therefore, always contain the same energy. He carries out a Lorentz transformation to the moving frame to obtain the surface;

$$\beta^2(x'+vt'-l'ct'-l'x'v/c)^2+(y'-\beta m'ct'-\beta m'x'v/c)^2+$$
$$(z'-\beta n'ct'-\beta n'x'v/c)^2 = r^2 \qquad (4)$$

where (l', m', n') are the direction cosiness as seen in the moving frame and

$$\beta = \frac{1}{\sqrt{1-v^2/c^2}}$$

where v is the relative velocity of the reference frames.

The frames are considered to coincide for $t = t' = 0$. To compare the observed energies, Einstein considers the ellipsoidal volume contained by (4) when $t' = 0$ and assumes this to be equivalent to the spherical volume contained by (1) at $t = 0$.

Once more the same mistake has been made. Points on (4) are not regarded by the moving observer as occurring at the same time. However, it should be noted that the quotations (i) and (ii) are not really inconsistent, as (i) refers to a spherical wave-front of light, while (ii) refers only to a spherical volume of space moving with the light beam.

Having had the temerity to couple Einstein and Keswani together and suggest they are both mistaken, does this mean Special Relativity is inconsistent? I think not. Where is the mistake?

In his derivation of the Lorentz transformation Keswani makes the unnecessary assumption that a spherical wave-front would transform into another spherical wave-front for the moving observer and then finds the assumption inconsistent with his conclusions.

In fact, only the constancy of the velocity of light need be assumed to obtain the relation,

$$x^2+y^2+z^2-c^2t^2 = 0 = x'^2+y'^2+z'^2-c^2t'^2$$

between the space-time *points* (x', y', z', t') as seen in S' and (x, y, z, t) in S at which a particular event occurs.

From this, we can derived the transformation between the *same event* as seen in different frames. We need not consider an entire spherical surface.

Einstein used only the assumption of the constancy of the velocity of light in his derivation of the Lorentz transformations. Hence, the fact that a spherical wave-front is not transformed into another spherical wave-front by a Lorentz transformation is not inconsistent with Einstein's basic postulate.

---

[1] *The Principle of Relativity* (Dover), A. Einstein. 'On the Electrodynamics of Moving Bodies', p. 57

In his second paper,[1] Keswani describes an experiment which he regards as a counter-example to the possibility of a Principle of Relativity. A spherical wave starting at O, the origin of a rectangular system of axis, passes through a green filter placed on the y-axis and the light is therefore seen as green at $P(0, r, 0)$.

In S', P is now observed at be at $P'(-rv/\sqrt{c^2-v^2}, r, 0)$. This at least is a step forward from his first paper, as Keswani now realises that such a point P' is no longer regarded as being on the y'-axis of S'. He maintains that white light is seen at P', the filter still being on the y'-axis of S'. It is difficult to see why, if a point P' is displaced from the y'-axis the filter which is also on the same axis should not also be displaced so that green light is seen at P'.

Keswani called Einstein the prophet. It seems we must thank Mr Keswani for showing us that even prophets make mistakes.

H. B. LEVINSON

MAN, BEAST, AND PHILOSOPHICAL PSYCHOLOGY

In their criticism of *Mental Acts*, recently published in this *Journal*,[2] Professors John King-Farlow and Elton A. Hall allege that its author 'seems to make little secret of his belief that men have souls and brutes do not' (p. 100). I hold no such belief, and the pages of *Mental Acts* cited as evidence by the authors (pp. 17, 121) will be found to contain nothing even resembling an expression of such a belief. (Perhaps 'souls' here was short for 'eternal souls', as in the preceding italicised sentence; but I do not believe *any* soul ought to be called eternal.) The whole passage looks like an appeal to prejudice—cf. the remark about 'religious scruples'.

PETER GEACH

---

[1] *The British Journal for the Philosophy of Science*, 1961–62, **12**, 151

[2] XVI (1965), 81

REVIEWS

MANDELBAUM'S CRITICAL REALISM

THIS is a most welcome book.[1] It deals soberly and realistically with important historical and philosophical issues. Its well-documented historical sections are motivated by philosophical questions, and its philosophical outlook is strongly influenced by the author's appreciation of some of the historical relations between science and philosophy.

The book consists of four related studies : on Locke; on Newton and Boyle; a criticism of Hume; and a defence of what Mandelbaum calls a *radical critical realism*. Since the whole book is informed by the philosophical view explicitly defended in the last study, I will take this first.

By *realism* Mandelbaum means the claim that a physical world exists independently of experiences of it, and that, moreover, we can learn something about this physical world. By *critical* realism he means, in addition, ' that at least some of the characteristics which physical objects appear to possess are not actually possessed by them '. By *radical* critical realism he means, further, ' that we do not have the right to identify *any* of the qualities of objects as they are directly experienced by us with the properties of objects as they exist in the physical world independently of us ' (p. 221); in other words, there is no limit to the extent to which commonsense perceptual appraisals of things may need to be modified in the light both of further empirical observations and of theoretical scientific considerations : whether or not our senses systematically deceive us, they do systematically under-inform us about the real character of the things to whose existence they alert us. And, of course, there are ever so many things of whose existence they give us no indication.

All this is agreeable to the present reviewer (to whom an idealist is like a prisoner who tries to comfort himself by pretending that there is nothing beyond the walls of his prison). Also agreeable is Mandelbaum's cheerful disregard of the various *cordons sanitaires* which a good many philosophers (with some encouragement from theologians and scientists) have tried to establish around certain main areas of thought (especially when new ideas in one area would otherwise have disturbing implications for old ideas in another area). Such *cordons* were put around science to stop Copernicanism, Darwinism, etc., causing trouble on the religious side, and also to stop microphysics from disturbing our everyday conceptions of things. (Many contemporary philosophers still assume these *cordons* to be effective. Thus Ryle has declared, *Dilemmas*, p. 81, that apparent conflicts ' between science and theology or between fundamental physics and common knowledge' are mere 'seeming feuds'.) The *cordon* between science and common sense tended, in turn, to insulate *epistemology* from science. For if the fundamental concern of epistemology is supposed to be the validity of our more primitive bits of knowledge ('Do I *know* that I am holding a pen?' etc.), and if physics is considered irrelevant to such knowledge, then physics will be considered irrelevant to *serious* epistemology, though epistemologists in holiday mood may take a look at the validity of theoretical scientific knowledge. (In view of this extending system of linked *cordons*, it is not surprising that philosophy came to be regarded as so insulated that *it* cannot disturb *anything* outside its own shrinking province.)

---

[1] *Philosophy, Science and Sense-Perception: Historical and Critical Studies.* By Maurice Mandelbaum. Johns Hopkins Press, Baltimore, 1964. Pp. ix+262. 52s.

Mandelbaum clearly regards *cordons sanitaires* as bad for philosophical health. He rejects (pp. 196, 238) the view that theoretical physics and common sense are essentially complementary and non-competitive; and he consequently rejects the view that science is irrelevant to serious epistemology. (Indeed, the view that the area of man's greatest cognitive achievements is irrelevant to a theory of human cognition is really quite preposterous.) Mandelbaum states in his Preface that critical realism is characterised by

> the belief that scientific inquiries are directly relevant to epistemological issues. And, in my opinion, it is in no small measure due to the fact that contemporary philosophers tend to draw a sharp distinction between scientific and philosophic problems that critical realism has suffered an almost total eclipse in our time (p. vii).

With regard to Locke, Mandelbaum advances three main theses: (1) Locke was a critical realist (in Mandelbaum's sense); (2) he was, moreover, a physical atomist; (3) his realism and atomism were consistent with his empiricist theory about the origin of ideas.

Another way of putting (1) is to say, as Mandelbaum does, that Locke should be viewed, not so much as a precursor of Berkeley and Hume, but rather as a philosophical ally of Boyle and Newton. I think that Mandelbaum is completely successful, here. Indeed, he goes some way towards establishing that Locke was a *radical* critical realist. For he shows that Locke's statement that our 'ideas of primary qualities of bodies are resemblances of them, and their patterns do really exist in the bodies themselves' means, not that our ideas of a body's primary qualities resemble those qualities *exactly* or even adequately, but only that our ideas of a body's bulk, figure, motion, etc., do not *radically* misrepresent its physical qualities (in contrast to our ideas of its colour, taste, smell, etc., which *do* radically misrepresent those of the body's qualities which cause us to have these ideas).

As to (2): I do not think that Mandelbaum establishes that Locke was a full-blooded atomist. On pages 4-5 Mandelbaum quotes or refers to eleven passages which, he claims, 'illustrate Locke's unquestioning acceptance of atomism'. Six of these passages seem to me to offer very little support for Mandelbaum's thesis; the other five show that Locke did indeed postulate what he variously termed 'insensible corpuscles', 'minute particles', or 'insensible particles'. But seventeenth-century atomism involved more than this: it was the existence of *indivisible* corpuscles that was denied by the anti-atomists, Descartes and Leibniz, and affirmed by Newton. (Mandelbaum quotes Newton's supposition that 'God in the beginning' formed particles 'so very hard, as never to wear or break in pieces'. To the Leibnizian question, but what holds together the constituent *parts* of an atom? Newton's answer was: no power can 'divide what God himself made one in the first Creation'.)

Now in none of those eleven passages did Locke affirm the existence of *indivisible* corpuscles; and how could he have, in view of his emphatic declaration (*Essay*, IV, iii, 25) that 'our want of precise distinct ideas of their primary qualities keeps us in an incurable ignorance about [these insensible corpuscles]'. No doubt, Locke was well-disposed towards the atomism of Boyle and Newton: but I think that Mandelbaum goes too far when he says that Locke regarded atomism 'as an empirically based conclusion drawn from the experimental inquiries of his day' (p. 60).

As to (3): I have the impression that Mandelbaum's confidence in the reconcilability of Locke's physical realism with his empiricist epistemology gradually waned as his study of Locke proceeded. On page 16 Mandelbaum roundly declares that he will 'now attempt to show' that they can be reconciled. The attempt is then postponed, however, and it is still being postponed three-quarters of the way through the essay (on p. 46 Mandelbaum says that before answering 'the fundamental question of this essay—in what measure Locke may be said to have been consistent in his realism', he will first attend to something else). When, on the last page, he finally gives his answer to his 'fundamental question' it is pretty muted:

> I should not wish to claim that Locke's reply to the challenges which came to be posed by Berkeley and Hume are adequate. However, it cannot be denied that Locke did see the problem, in outline at least, and did attempt to propose a solution to it (p. 60).

The theories of modern science transcend empirical evidence not only horizontally but also vertically: they are not only universal in scope but postulate entities beyond the reach of human observation. Mandelbaum writes as an inductivist, as someone who holds that the gaps between evidence and theory must be bridged by non-demonstrative or inductive inferences of some kind. He calls a vertical 'inference' from empirical evidence to unobservable entity *transdiction* (a term he borrows from D. C. Williams), and he regards the 'problem of transdiction', i.e. the problem of justifying such vertical 'inferences', as a major part of the problem of induction.

Mandelbaum disregards the alternative view, associated primarily with Popper, whereby there is no inductive inference at all from evidence to theory, and hence no vertical or transdictive inference, and hence no problem of justifying such inferences. But in connection with Mandelbaum's historical study of Newton and Boyle, this disregard is actually an advantage. A Popperian historian might wave aside their attempts to present their corpuscular theories as somehow derived from the phenomena as so much unnecessary, and necessarily unsuccessful, inductivist apologetics. But Mandelbaum's inductivism obliges him to examine their attempts sympathetically; and he emerges with, among others, the interesting result that the Third of Newton's 'Rules of Reasoning in Philosophy', as formulated in the second edition of the *Principia*, was intended precisely to entitle him to make 'transdictive' inferences from observed characteristics of medium-sized bodies to conjectured characteristics of corpuscles.

I now turn to Mandelbaum's criticisms of Hume's sceptical arguments. I will single out for attention one particularly interesting criticism, which might be summarily re-stated thus.

Hume argued that we do not really have some of the ideas we are commonly supposed by realist philosophers to have, because all our ideas are derived from antecedent impressions, and no idea could be so derived of, say, the continuous existence, independent of experience, of a material object, or of the correspondence or otherwise between the real physical characteristics of a material object and an observer's impression of that object. For Hume's argument to work it was, of course, essential that he should be able to *distinguish* the primary elements (impressions) from the derivative ones (ideas). If he could not have done this, he would have had no answer to anyone who roundly declared: 'But I *do* have an idea of the independent existence of a material object'; for then, Hume could no longer have claimed that the primary elements do not permit the derivation of any such idea.

Hume himself conceded that his attempt to distinguish between impressions and ideas in terms of their relative vivacity does not always work satisfactorily. If we reject any such purely subjective or psychologistic criterion, as surely we must, what remains? Mandelbaum asserts, I think rightly, that what Hume really relied on, in making this crucial distinction, was the assumption that impressions, unlike ideas, are directly caused by the action of the external world on our sense-organs. Let us use the letter *k* to designate the realist assumption that a physical world exists externally to us and causally influences our perceptions of it. Then we could summarise Mandelbaum's main criticism of Hume thus: Hume presupposed *k* in his argument against *k*. Mandelbaum is careful not to accuse Hume of a formal inconsistency, here—he says that Hume's argument presupposed *k*, not in the sense that it entails *k*, but in the sense that, without *k*, there would be no way of securing the distinction on which the argument essentially depends.

Mandelbaum's book ends a bit limply; and in some places his discussion might, perhaps, have moved at a quicker pace. But such minor aesthetic shortcomings are the obverse of the book's patent intellectual decency.

J. W. N. Watkins

Language and Information—Selected Essays on Their Theory and Application
By Y. Bar-Hillel.
Addison Wesley, London and Jerusalem Academic Press, 1964, pp. x+388. $12.50.

THE author of this useful collection of reprints delights in the role of gadfly, even where he finds it necessary to castigate himself for youthful enthusiasms now abandoned. 'A close study of one of the most publicised schemes for the mechanization of problem solving, and a somewhat less detailed study of the whole field of artificial intelligence, has shown an amount of careless and irresponsible talk which is nothing short of appalling, and sometimes close to lunatic. There is absolutely nothing in all this talk which shows any promise to be of real help in mechanizing translation.' With such purple passages, this book is certainly not lacking in human interest, particularly as Dr Bar-Hillel confesses himself one of the prime movers of the great bandwaggon of Mechanical Translation.

The snag in 'M.T.', of course, arises from the part played in human translation by background knowledge of the *subject-matter* of the passages to be translated. The last word in the sentence 'The box was in the pen' demands (for us) to be translated by a word signifying an enclosure rather than a writing implement, because we know that the physical situation described is more probable in this sense than in the other. Bar-Hillel sees no prospect of a computer's being equipped (except *ad hoc*) to make a comparable judgment. His reasons for scepticism are perhaps the most philosophically interesting feature of the book, which also contains papers on algebraic linguistics, information theory and the mechanization of information retrieval. The refreshingly candid introduction will be found worth reading before any guidance is sought in the remainder.

How much truth lies in Bar-Hillel's 'statement of impotence'? In one sense, it is clearly true and important, and constitutes a valid criticism of a large bulk of recent claims unilluminated by its insight. A computer translation programme restricted to purely formal clues can never be guaranteed against being fooled by an expression whose formal ambiguity is easily resolvable (and therefore tolerable in practice) by sufficiently knowledgeable readers who understand what it is about; moreover 'the distance from symbol manipulation to linguistic understanding is enormous, and loose talk will not diminish it' (p. 214).

No one would wish to minimize the difficulties in equipping a computer programme with the 'faculté de langage' that Bar-Hillel depicts as the end-product of 'hundreds of millions of years of evolution'. Yet since the arithmetical faculty, in which the computers excel, is itself at least equally ancient in lineage, it is hard to feel that this evolutionary argument carries much weight. The deficiencies of most computer models of linguistic comprehension, I believe, have nothing to do with their lack of genetically determined 'special purpose' structure (p. 215). They arise in principle simply from the lack of an adequate internal *analogue* (in a general sense) of the domain of discourse.[1, 2]

This lack could be remedied at two levels. First, the computer could be equipped with an internal 'workshop', whereby appropriate skeleton analogues of rival interpretations of an ambiguous sentence could be constructed (with the help of an 'operational dictionary' or its 'learned' equivalent), and the outcome of this internal experiment used to evaluate the relative probability of each interpretation. Thus 'the box is in the pen' could be automatically turned into a programme for the representation by a skeleton analogue of a 'box' enclosed by a 'pen'. The outcome (possible/impossible/improbable) would depend on the nature and dimensions of the analogues which past evidence (or operational definition) had led to be coupled with these words. If 'impossible' with a given analogue, this finding could automatically generate a search for an alternative analogue for each of the operative words.

The performance of such a device might well rival that of a human translator in any domain in which adequate operational instruction had been given; but it could still be questioned[3] whether it exhibited a genuine linguistic understanding of a linguistic expression—i.e. the ability to perceive it as a *tool* constructed by a cognitive purposive agent and designed to serve a specific purpose in its impact upon a similar agent. For linguistic understanding at this level, the internal analogue-making facilities of the computer would have to be extended in ways that cannot be discussed here[4]; but it is not suggested that these would be essential in a satisfactory mechanical translator. The purpose of the foregoing is only to rebut any suggestion that the semantic deficiencies of present day computers *vis-à-vis* human brains have no presently specifiable remedy. True, no 'box of wires and valves' could ever validly be said to understand anything; but then, no 'mass of brain cells' does so either.[5] If we carefully compare like with like, it is hard to see why the semantic capacities of human brains should not be adequately paralleled in those of human artefacts, even in their present state of development.

The section on information theory, written more than ten years ago, is lucid and still valuable, though naturally the references to other works are

---

[1] K. J. W. Craik, *The Nature of Explanation*, Cambridge University Press, 1943

[2] D. M. MacKay, 'Mindlike Behaviour in Artefacts', *Brit. J. Phil. of Sci.*, II, 105-121, 1951; also III, 352-353, 1953

[3] D. M. MacKay, 'Linguistic and Non-Linguistic Understanding of Linguistic Tokens' RAND Memo. RM-3892-PR, February 1964

[4] See for example, M. Kochen, D. M. MacKay, M. E. Maron, M. Scriven and L. Uhr, 'Computers and Comprehension' RAND Memo. RM-4065-PR, April 1964

[5] D. M. MacKay, 'The Use of Behavioural Language to Refer to Mechanical Processes', *Brit. J. Phil. of Sci.*, *XIII*, 89-103, 1962.

no longer representative. Its warnings are still in place, alas, against confusing the concept of information *per se* with a statistical measure of the rarity of the symbols in which information is expressed. On the other hand it may be doubted whether the author's (1952) attempt with Carnap to base a theory of semantic information on inductive probability (chapter 15) can do much more than quantify one important aspect of semantic information—namely, its power to reduce uncertainty. A theory to do justice to all semantic aspects of information must, one feels, demand closer attention to psychological as well as logical considerations. But we need not wait for this before beginning upon the ' terminological cleanup ' or ' debabelization ' (*sic.*, p. 296) whose necessity is so pungently argued here.

D. M. MACKAY

The Act of Creation. By Arthur Koestler
Hutchinson, London, 1964. Pp. 751. 42s.

THIS book is not about God. So it is not about creation in the theological sense of making something out of nothing. It is about men getting new kinds of idea or making new kinds of sentence or thing. Mr Koestler's ' creation ' is thus by definition the production of something *of a novel kind.* It is usually also the production of something of a novel *and good* kind; getting a novel idea tends not to be ' creation ' for Koestler if it is a worthless idea.

Koestler offers us some 634 pages about this occurrence of getting a novel and good idea, plus another 100 pages of appendices, notes, booklist, index, and foreword by Professor Sir Cyril Burt. How on earth, one wonders, does he find so much to say about it? The wonder grows when we observe that he gives little space to the most obvious topic, namely what dodges there are for increasing the frequency with which one gets a new and good idea. This book is not a rival to Polya's *How To Solve It*, though it quotes a little therefrom.

Koestler finds so much to say about ' the act of creation ' largely by regarding as relevant to his theme half the science of psychology and many other topics. We have here substantial reports of much of what has been written on laughter, smiling, emotion, humour, tickling, the mentality of apes, the unconscious, the history of science, the limits of confirmation in science, the place of emotion in science, weeping (there is a special bibliography of books on weeping), hierarchic structure in organisms, illusion in art, rhythm, rhyme, pictorial thinking, character, plot, the aesthetic experience, snobbery, receptivity in art, the work of deoxyribonucleic acid in the chromosomes, the growth of the embryo, the development of the nervous system, principles of organic organisation, codes of instinctive behaviour, imprinting and imitation, motivation, playing and pretending, perception and learning and memory (but 'the neurophysiological problems of memory are beyond the scope of this book', 519), motor skills, the controversy in psychology between Gestalters and associationists or stimulus-response men, learning to speak, and learning to think. The book is something like a digest of recent psychology with special emphasis on the getting of new ideas.

The main doctrine recommended is that the act of creation is a bisociation of two hitherto unassociated matrices belonging to different levels. One of the examples given is that Kepler got radically new ideas on planetary motion by bisociating the hitherto unassociated fields of astronomy and physics. A 'matrix' is a field or system of thoughts. 'Bisociation', a term invented by Koestler and in my opinion not properly defined, is meant to contrast with the common notion of association, but seems in use to mean just a new and surprising association.

I cannot think that it needed these 600 learned pages to arrive at and recommend this doctrine. I cannot think that most of them have any relevance to the doctrine. I cannot think that the doctrine is of any value. Having declared that Kepler got an original idea by bringing physics into astronomy, what does one add by saying: 'He bisociated two hitherto unassociated matrices belonging to different levels'? It is the same thing in worse language. It adds nothing to the existing stock of advice on how to be original. It explains nothing; for, if one mathematician is original and another is not, it is no explanation to say that this is because the second does not bisociate hitherto unassociated matrices. It predicts nothing; for, if Koestler wants to say that *every* creative idea will be found to involve bisociating two hitherto unassociated matrices, he will have to make any little difference count as two matrices. In fact, not all great discoveries involved going outside the science as then conceived. 'When Einstein bisociated energy and matter', as Koestler puts it (233), that was not going outside physics.

The sad phenomenon which we have here (and I have known another instance of it) is this: a successful creative artist begins to wonder how he did it, to philosophize on his creation: he assumes that an answer is to be found, and that it is to be found in some kind of learning or study or experiment; he embarks on a big course of study, and uncovers many interesting facts which he thinks or hopes are relevant; but, as he is not competent in analytical philosophy, he arrives at a conceptual explanation which, while it seems to him to embody both the results of his reading and his experience as a creator, is in fact a disguised tautology. And in the process, alas! Koestler the limpid and exciting storyteller beomes Koestler the obscure and loose-thinking student. 'Built-in perceptual analysers for the recognition of species-specific sign-releasers may be described as the result of the phylogenetic abstraction of biologically relevant stimulus-patterns. We may regard this abstractive power, with Lashley, as " one of the primitive basic functions of organized nervous tissue " ' (608). That a writer of Koestler's excellence should come to this!

Many traces of the good storyteller remain, thank heaven. All the old stories about famous discoveries are well retold, and many that I never heard before. Passages of unhelpful jargon are relieved by passages of sprightly though somewhat journalistic writing.

Koestler has published at least one good work of scholarship; for his *The Sleepwalkers* is a first-hand, illuminating, and novel study of the works of the great astronomers. The present book unfortunately is not like that; it is a second-hand and unilluminating digest of a great many other people's works, many of which are themselves digests.

We do not know how we get new ideas, and probably never shall know. It is hardly a joke to say: Why is this book so long? Because there is nothing to say about its subject.

RICHARD ROBINSON

ABSTRACTS

Dialectica, 1963, 17, Nos. 2/3

M. Maruyama, 'Basic Elements in Misunderstandings II'

Basic concepts of communicational elements are explained: culturally institutionalised or individually internalised conscious or unconscious, intentional and unintentional encoding and decoding systems of verbal and non-verbal behaviour, metacommunicational cues, various modes of communication (factual, insinuative, deceptive, etc.), the role of situational context and of background information.

Examples of complex misunderstandings in interpersonal interactions are given and analysed. This analytical method has been useful in psychiatry as well as in daily-life situations.

L'Analogie en tant que Methode de Connaissance

F. Gonseth, 'Extrait d'une lettre à MM. Cuénod et Pun'

M. Cuenod et L. Pun, 'Réponse à M. Gonseth'

F. Gonseth, 'Analogie et modèles mathématiques'

The author first investigates how to make precise and get to the core of the sense of a word in common language. Any definition, however appropriate it may seem, is in fact inadequate. The word must be considered as a moment of an activity of making precise, which is particularly clear in the case of analogy as a word and an idea. With that in view, the concept of the scheme is explained and it is examined in which horizon a mental scheme may have an existence of its own. In order to provide a satisfactory answer, it is necessary to investigate and uncover the structure of subjectivity, to make it understood that the basic principles of mathematics arise from elaborating some of these disciplines in relation with a scheme.

Two fields of existence may then be said analogous if they may be considered as two different realisations of a scheme. The views gained in investigating the structures of subjectivity permit us to realise (1) how the result may be extended to mathematical models and (2) how the part played by those mathematical models may be made precise.

J. Loeb, 'Par quel mécanisme l'analogie est-elle un source de connaissance?'

It is shown how, and for what reasons, analogy is a source of knowledge: first, it permits a less developed technique in a given domain to benefit from the progress of a more advanced technique. One can cite as examples: the pair hydraulics-electricity and the pair heat-electricity; in the latter case, analogy permitted creation of the notion 'heat resistance'.

Finally it is shown that analogy, like mathematics itself, creates information because it is abstraction, in the etymological sense of the word.

ABSTRACTS

E. Mesarovic et M. D. Henning, ' Analogy in the Creative Processes and the Objects of Creation in Art and Sciences '

The method of analogy is considered as it is used in epistemological considerations of both the object of creation and the creative process itself in art and the sciences. Both areas of creativity are considered within the context of the general systems theory concept of an open system which offers a convenient vehicle for relating the scientist and the artist with the product of his creation. It also provides a convenient method for the explanation of the essential nature of the created object. It is shown how the creative process in both art and science can be viewed as a heuristic search for the structure of an open system—the object of creation.

Similarities in the two areas—art and science—are emphasised over the differences, in order to demonstrate the power of the method of analogy.

G. Pask, ' The Use of Analogy and Parable in Cybernetics with Emphasis upon Analogies for Learning and Creativity '

F. H. Raymond, ' Analogy et intelligence artificielle '

P. Naslin, ' Analogies, homologies et modèles '

After a few considerations on the limited value of reasoning by analogy, the author distinguishes between analogy, homology and model, as illustrated by electro-pneumatic, electro-mechanical and electro-thermic homologies. On the other hand, there exist numerous models of a same process: mathematical and symbolic models, graphic models, analogic and numeric models. The paper ends with remarks on the limits and the applications of models, namely in adaptive systems, as well as on their philosophical scope.

L. Pun, ' Optimum de l'analogie '

Progress in automatic control, particularly in optimising and self-adaptive controls, is used to establish models in order to connect three horizons: knowledge, human activities and automatic controls. The guiding operation in all these horizons is recognised as evolution. In the horizon of automatic controls, the evolutions can be said closed, since they are completely defined. In the horizon of human activities, they can be said to be open, because the overall direction depends both on the constituents and on the partial accomplishment. In the horizon of the acquisition of the knowledge, there appear to be both evolutions in closed and open form; the problem is therefore to know whether these evolutions are teleological or homeostatic. There does not seem to be a ready answer; it could however probably be accelerated by better knowing the concept of matching the analogy.

U. Zelbstein, ' Méthodes analogiques et isomorphisme des chaînes fonctionnelles '

The knowledge of an object under observation will have no concrete meaning unless we can work out a logical model the similarities of which are known in relation to the analysed structures. The logical model, when dealing with complex structures, is drawn out by similarity, starting from an analogical model worked out in relating local dimensions in the object to physical dimensions of a different nature which restrict the number of variables and can be arranged and modified as required.

Any operating similarity and behaviour comparisons come actually down to an apparent functional analogy. The difference between the thin structure of the ' natural ' and artificial systems makes it illusive to hope that a strict isomorphism between these two fields may thus be sorted out.

However, the fact of achieving analogical models widens our knowledge and shows new possibilities to the human creative mind.

M. Cuenod, ' Analogie et notion de rigueur pour l'ingénieur '

Analogy plays a basic part in engineering which is subject to certain limitations, due to the following reasons:

(1) lack of accuracy of the measuring instruments and of computation data;
(2) simplifying assumptions necessary for establishing the mathematic model.
(3) restrictions due to compulsory computation means and time.

The basic criterion of engineering knowledge is efficiency. The engineer works thus on successive approximations; each of them corresponding to a more or less appropriate model of the final result. In order to avoid disappointment, it is very important for him to be conscious of the relativity of his knowledge and of the scope of the successive approximations.

This procedure is the same for other fields of knowledge and provides an opportunity to specify the concept of rigour applying to each. It shows the practical, limited and revisable character of any knowledge not springing from metaphysics.

M. Hesse, ' Analogy and Confirmation Theory '

Applications of analogy in science can be grouped under three headings:

(1) *Formal analogy*, in virtue of mathematical isomorphism;
(2) *Heuristic analogy*, when a relation of similarity between the object to be explained and a model is used to suggest an explanatory theory;
(3) *Inductive analogy*, when the model is *identified* with the theory in such a way that reasoning by analogy with a model is relied upon to give predictions in the domain of the object.

There can be no justification for using analogies in the way described under (3) rather than (1) or (2) unless we have an inductive logic of analogy, and not even Carnap's confirmation theory in its first version is strong enough for this. His latest theory does, however, include an explication of analogical argument.

Dialectica, 1963, 17, No. 4

J. Edeline, ' Aspect Logique du Jeu de Mots Poetique '

Certain poets tell stories, others only appreciate the ' palpable aspects of signs ' (Jakobson). The latter frequently use puns, thus stressing *formal analogies* between words (rhyme, assonance, etc.).

This tends to make us accept that formal analogy results in sense analogy. The logical structure of this reasoning (implicit with poets) is explained and shown erroneous, for, contrary to science, metaphors and puns in poetry are equations or models *which do not operate.*

But these proceedings are in use and it is inadequate to claim that they are but deceits or illusions. In fact poetry is strictly imaginary, its universe is one of pure representation and it refuses any checking by reality or experience—which is of basic importance in any scientific proceeding. In the end, this characteristic appears as the only real difference between poets and scientists, both of whom work by application of analogical models.

J. Zafiropulo, 'Le rôle de l'analogie dans le déchiffrement de l'écriture mycénienne linéaire B'

L. Solari, 'Contenu et portée des modèles économétriques'

The article studies the logical and material content of econometric models, which constitute the chief instrument of knowledge and action in economic science. It is based on a general formalisation which makes the essential components of models explicit, and which covers most actual applications. Analysis of conceptual and concrete problems raised by the specification of econometric models leads the author to study their bearing on two principle uses, often intimately connected: prediction and decision. The need to integrate better the theoretical and experimental approaches reveals itself in the discussion of difficulties in applied economics, and of the need to obtain operational models.

G. E. von Grunebaum, 'Muslim World View and Muslim Science'

C. T. K. Chari, 'Information, Cosmology and Time'

In this paper the methodological and philosophical implications of modern information theory for cosmology, especially for the rôle of time in 'world-models', are discussed. The different approaches to measures of information and their bearing on cosmological 'timescales' are examined. Attention is paid to stochastic models, in particular Markovian and ergodic models, in relation to information theory. The question of the invariance of temporal structure for all changes of 'scale' is mooted. The suggestion is advanced that 'informational uncertainty' may be intrinsically bound up with all large-scale observations of the universe. Reference is made to the statistical models of J. Neyman, E. L. Scott and G. C. McVittie and the cosmological 'uncertainty' enunciated recently by W. H. McCrea.

Methodos, 1961, 13, Nos. 51/52

P. S. Ray, 'The Logic of Linguistics'

This is a critical review of A. Noam Chomsky's work in theoretical linguistics, in so far as it bears upon a philosophy of language. Three possible interpretations of his position are examined and rejected. The more narrowly technical linguistic aspect of his work is positively, but more briefly, evaluated.

P. S. Wadia, 'Ethics and the "Ought"-premise'

The central theme of the paper is a critique of R. M. Hare's *The Language of Morals*. It would seem that in Hare's meta-ethical formulations an unsuccessful attempt is made to weld together a purely psychological doctrine concerning the decisional factor in choice situations (the factor of 'commitment' as the existentialists like to call it) with another doctrine (of a vaguely utilitarian kind) which seeks the justification of moral judgments in the effects of observing them. But one cannot have it both ways: Either the answer to the question ' why *ought* one to do this rather than that? ' is chiefly in terms of the effects expected to follow from actions, in which case the doctrine is naturalistic, or in terms of decision-making, in which case morality becomes an arbitrary matter.

M. Maruyama, ' Awareness and Unawareness of Misunderstandings '

The article gives empirical examples from daily life situations and from cultural contact situations for the theoretical framework previously published in *J.P.S.*, 1961, 12, pp. 52-62. Examples are taken from anthropological (ethnological) materials, and from confusions in cultural situations such as the Cargo Cult.

T. Kotarbinski, I ' The Aspirations of Praxiologists '
II ' Postulates for Economic Modes of Action '
III ' The Property of a Good Plan '

The three articles together constitute an introduction to praxiology. In the first it is claimed that praxiological issues form the issues of the so-called theory of organisation and management. There follows a summary of the book *A Treatise on Good Work*. The second article deals more fully with one of the topics from this book. This is a review of the main forms of economic activity, those forms through which disposable energy is economised or used in a more profitable way. The last article reviews the values with which a good plan must be imbued. It should be purposeful, performable, coherent, flexible, operative and rational. In connection with the postulates of performability and rationality attention is drawn to their fundamental disharmony, since fully rational activity should be prefaced by proofs of performability of an activity, while final proof of performability of an activity is its actual execution.

Methodos, 1962, 14, No. 53

F. R. Harrison, ' Wittgenstein and the Doctrine of Identical Minimal Meaning '

The ' Doctrine of Identical Minimal Meaning ' purports to tell us how and why we can and do meaningfully use our ordinary language. In brief, the doctrine states that every general word such as ' good', ' blue ', ' tool ', ' game ', and so on must have one essential meaning to which it refers or names in all of its varying uses. Historically, I have chosen to discuss in this paper two men who have held such a doctrine; viz. Thomas de Vio Cardinal Cajetan and A. J. Ayer. And, drawing heavily on the notion of family resemblances found in Wittgenstein's *Philosophical Investigations*, I have attempted to dispel the hold this doctrine of identical minimal meaning has over us, by an examination of how we actually come to learn how to use our language.

F. R. Berger, ' Stevenson, Meaning, and Metaethics '

The theory of meaning presented by Charles L. Stevenson in *Ethics and Language* is considered for the light it can shed on the relations between metaethics, so-called theories of meaning and semiotic or the general theory of signs. Proposals are made concerning the use of the expressions ' theory of meaning ' and ' theory of signs '. It is argued that since ' meaning ' has had such divergent technical and non-technical uses, any theory of meaning is, at least in part, prescriptive. A general theory of signs is taken to be a metalanguage for talking about *all* aspects of a meaning-situation (not just those some ' theory of meaning ' singles out as *the* meaning). The relations of these respective kinds of theories to metaethics are explored and a brief metaethical view is presented.

K. Nielsen, ' On Looking Back at the Emotive Theory '

The emotive theory is re-examined so as to bring to light both its crucial insights and defects. Moral utterances are shown to be neither empirical statements of fact nor statements of ' moral facts ' known by some faculty of moral intuition. But while moral claims are not verifiable or capable of demonstration, we do give reasons for such claims. It is argued that the emotive theory's concept of good reasons is inadequate and that it also fails to note that moral discourse is a form of practical discourse. An alternative account of moral reasoning is sketched that overcomes these defects while still preserving the emotive theory's insight into the close connection between our moral beliefs and our attitudes.

Methodos, 1962, 14, No. 54

K. M. Sayre, ' Human and Mechanical Recognition '

Humans commonly recognise patterns and objects without reference to specific criteria, even in cases where critria could be specified if desired. In learning to recognise patterns, moreover, humans are capable of isolating invariant features of objects which are unanalysable, or are not analysed, in terms of other features which previously they had been able to isolate perceptually. After a review of currently prominent techniques for the mechanical recognition of patterns, it is argued that mechanical recognition systems are incapable (1) of identifying objects the invariant features of which have not been analysed, and (2) of generating, through adaptive programming or any similar method, an invariant feature which is not a combination of features to which the system has been explicitly prepared to respond. This suggests that the significant differences between human and mechanical recognition abilities lies more in the area of data acquisition than in that of data processing. It is the former area that a ' breakthrough ' is needed if machines are to approximate human recognitive behaviour.

J. Margolis, ' Intention, Consciousness, and Action '

Seven theses of Stuart Hampshire's regarding the concept of intention are examined. In opposition to Hampshire, it is argued that there is no necessary connection between consciousness and intention, or between action and intention, or between consciousness and action. Hampshire is specifically criticised for failing to distinguish adequately between intending to do something and predicting one's own behaviour and, generally, between the psychological and logical relations between the concepts of intention, consciousness and action. Attention is drawn to the difficulty of knowing what kinds of counter-arguments Hampshire would acknowledge and to the way in which Hampshire attenuates the concept of intention. Ordinary counterinstances are provided for all of his major theses, though he sometimes defines the categories in question in a way in which counter-instances would seem to be inappropriate. On the other hand, the argument assumes that Hampshire is offering a philosophical description.

M. Maruyama, ' Generating Complex Patterns by Means of Simple Rules of Interaction '

This article gives a concrete model for increasing the amount of non-redundant structuredness, and hence the amount of non-redundant information, by means of deviation-amplifying mutual causal relationship within an informationally closed deterministic system. The implication of this model is as follows: the amount of information contained in the embryo may be far smaller than the amount of information needed to describe the adult structure; it may be impossible to infer from the adult structure the embryonic information regarding the rules of pattern generation; two identical embryos under such a model always generate two identical adults.

The Methods of Contemporary Thought

by J. M. BOCHEŃSKI

(Translated from the German)

1965, X + 135 pp., ƒ 18.— (36s. $5.05)

Contents Prologue — Introduction — The Phenomenological Method — Semiotic Methods — The Axiomatic Method — Reductive Methods — Epilogue

Mathematical Thought

An Introduction to the Philosophy of Mathematics

by E. W. BETH

(Synthese Library)

1965, XII + 206 pp., ƒ 30.— (60s. $8.40)

Contents Introduction — Criticism of Mathematics as Based on Naive Intuition — The Foundations of Arithmetic — Symbolic Logic and its Connections with Traditional Logic — Intuitionism and Formalism — The Paradoxes — Significs and Logic — Recent Developments — Concluding Remarks — Bibliography — Index of Subjects — Index of Names

D. REIDEL PUBLISHING COMPANY / DORDRECHT-HOLLAND

DIALECTICA

VOL. 18 No. 1-4 15.3 - 15.6. 1964

Contents	*Sommaire*
	LES ANTINOMIES EN DROIT
PERELMAN (Ch.)	Avant-propos
FORIERS (P.)	Les antinomies en droit
MALGAUD (W.)	Les antinomies en droit; à propos de l'étude de G. Gavazzi
MORGENTHAL (L.)	Les antinomies en droit social
VANDER ELST (R.)	Antinomies en droit international privé
SILANCE (L.)	Quelques exemples d'antinomies et essai de classement
HOEFFLER (J.)	Les antinomies en droit public
BOLAND (G.)	Quelques propos sur les antinomies et pseudo-antinomies, en particulier en droit administratif
HUBERLANT (Ch.)	Antinomies et recours aux principes généraux
BOBBIO (N.)	Des critères pour résoudre les antinomies
MIEDZIANOGORA (Ch.)	L'adage « Nemo auditur propriam turpitudinem allegans » et le Code Napoléon
SALMON (J.)	Les antinomies en droit international public
TAMMELO (I.)	Tensions and Tenebrae in Treaty Interpretation: « Reconnaissance of a Battlefield for « Antinomies »
SZABÓ (I.)	Des contradictions entre le droit des différents systèmes sociaux
BUCH (H.)	Conceptions dialectiques des antinomies juridiques
PERELMAN (Ch.)	Les antinomies en droit. Essai de synthèse.

EDITIONS DU GRIFFON NEUCHATEL SUISSE

JOURNAL OF THE HISTORY OF IDEAS

A Quarterly devoted to Intellectual History

VOLUME XXVII JAN.-MARCH 1966 NUMBER 1

TABLE OF CONTENTS

ARTICLES

Philosophy and Scientific Methods . . . . *Richard McKeon*
Buridan and Seneca . . . . *James J. Walsh*
Locke on Personal Identity . . . . *Henry E. Allison*
Problem of Gulliver's Fourth Voyage . . . . . . *John J. McManmon*
Idea of Westernizing Russia, 1830-1850 . . . . . . *Raymond T. McNally*
Malthus and the Idea of Progress . . . . *Samuel M. Levin*
The Emergence of Progressive History . . . . . . *Charles Crowe*

NOTES

Heraclitus and Fire . . . . *Constance I. Smith*
T. Hariot: Early Atomism in England . . . . . . *Robert Kargon*
J. G. Hamann: Metacritic of Kant . . . . . . *W. M. Alexander*
Swift's Flying Island . . . . *Robert C. Merton*
" Begging the Question " . . . . *A. W. Sparkes*

REVIEW

Nakamura's *Ways of Thinking of Eastern Peoples*
reviewed by *A. T. Embree, C. M. Schirokauer, A. E. Tiedemann*

Subscription Rates (outside U.S.A. post paid): $6.50 for one year;
$1.75 (plus postage) for single issue

Payable to

JOURNAL OF THE HISTORY OF IDEAS, INC.
The City College, New York, N.Y. 10031

Philosophy and Phenomenological Research

AN INTERNATIONAL QUARTERLY

Edited by MARVIN FARBER

SEPTEMBER 1965, VOL. XXVI, No. 1

CONTENTS

I. ARTICLES
Intrinsic Value, by MONROE C. BEARDSLEY
Old Mysticism and New Knowledge, by HORACE M. KALLEN
The Mote in the Eye of the Critic of Critical Realism, by W. PRESTON WARREN
Existentialism and Existentialisms, by MIKEL DUFRENNE
The Scale Step, by GERHARD ALBERSHEIM
The Cartesian Circle, by LYNN E. ROSE

II. DISCUSSIONS
On Rose's "Cartesian Circle", by NORMAN KRETZMANN
Brentano on the History of Greek Philosophy, by HUGO BERGMAN
Husserl and Pre-Reflexive Constitution, by RICHARD T. MURPHY, S.J.
The Problem of Pain and Contextual Implication, by THOMAS A. LONG
Lamont on Rights and Duties, by MARCUS G. SINGER
Toward Understanding Aristotle's Categories, by MICHAEL NOVAK

III. REVIEWS, NOTES AND NEWS, AND RECENT PUBLICATIONS

Subscription prices $6.00 per year for individuals; $7.50 per year for institutions. Separate issues, $1.65 and $2.00 respectively. Postage added for back issues only. A circular listing the contents of the Journal since 1940 will be sent upon request.

Address all communications to

PHILOSOPHY AND PHENOMENOLOGICAL RESEARCH
State University of New York at Buffalo, Buffalo, New York, 14214

Atlas Publ. & Distr. Co. Ltd., London - Stechert-Hafner, Inc., New York - Hermann & Cie., Paris - Nicola Zanichelli, Bologna - H. Bouvier u. Co., Bonn a/Rh. - Mezhdunarodnaja Kniga, Moscou - Gerold & Co., Wien - F. Machado & Cia, Porto - Libreria Editorial Augustinus, Madrid - Librairie Payot, S. A., Lausanne - The Maruzen Company, Tokyo.

"SCIENTIA"

INTERNATIONAL REVIEW OF SCIENTIFIC SYNTHESIS

A REVIEW COVERING ALL SCIENCES
(1965 - 59th year)

Scientific Committee: Chief-Editor : P. BONETTI
G. ABETTI - G. B. BONINO - P. CALDIROLA - L. CALIFANO - R. CANESTRARI - G. COLONNETTI - B. DE FINETTI - M. GORTANI - G. LEVI DELLA VIDA - G. MONTALENTI - E. PADOA - V. SOMENZI - S. TONZIG.

"SCIENTIA" is the only review of its type which: has a world wide circulation : deals with the most recent and fundamental problems of all branches of science : can boast among its contributors of the most illustrious men of science in the whole world : publishes the articles in the native language of their Authors (English, French, Italian, German, Spanish). Each issue includes a **Supplement** *containing the complete French translation of the articles which in the text are published in language other than French.*

"SCIENTIA" has therefore a very strong appeal to the scientific-minded reader all over the world.

Full details and a free back copy will be sent by applying to

"SCIENTIA" - ASSO (Como, Italy)

sending 1s. 11d., or equivalent amount in other currency, preferably in air-mail postage stamps of your country, merely to cover packing and postage. For a number of the current year, please send 20s. 5d., or equivalent amount in other currency, which will be deducted from the subscription price.

ANNUAL SUBSCRIPTION : £6. 15s. 5d.

PHILOSOPHIA NATURALIS

Archiv für Naturphilosophie und die philosophischen Grenzgebiete der exakten Wissenschaften und Wissenschaftsgebiete

CONTENTS OF VOL. VII/2

ARTICLES:

HUBNER: Was heisst und zu welchem Ende studiert man Naturphilosophie?
HEYDE: Das Bedeutungsverhältnis von φιλοσοφία und "Philosophie"
FREYTAG-LÖRINGHOFF: Zum Problem des Zufalls
MÜLLER: Gesetz und Denken in der exakten Naturwissenschaft
MÜLLER-MARKUS: Erkenntnisprobleme der Relativitätstheorie
SCHORR: Der Begründungszusammenhang der Konzeption Hugo Dinglers

Single copy DM 11 *; Subscription DM* 36 *per volume* (4 *issues*).

VERLAG ANTON HAIN KG . MEISENHEIM AM GLAN
(Deutschland)

The British Journal for the Philosophy of Science

VOLUME XVI FEBRUARY, 1966 No. 64

ON LOGICAL DEFINITIONS OF CONFIRMATION*

D. STOVE

I

CONFIRMATION, like other concepts of inductive logic, does not seem to be at all a psychological idea. There seems to be no reason why we may not regard it as an objective relation which holds between statements. If it is such, it is evidently not a relation the existence of which depends on the factual truth of the statements it relates. Furthermore, the empirical sciences, despite their different subject-matters, do not appear to employ correspondingly different concepts of confirmation: there seems to be one general concept which is common to them all.

Such considerations incline many philosophers to the view that it is possible (however difficult it may be) to characterise confirmation in purely logical terms, so that the confirmation of one statement by a second will be independent of the statements' truth and content in just the same way as is, for example, the entailment of one statement by a second.

Such a definition was proposed by Professor C. G. Hempel in a classical article.[1] So we may certainly number him among those who subscribe to the thesis:

1. Confirmation can be defined in purely logical terms.

Hempel begins his article, however, by asserting another thesis which is also widely held:

---

* Received 27.i.65

[1] 'A Purely Syntactical Definition of Confirmation', *Journal of Symbolic Logic*, 8, 1943. All my subsequent references, unless otherwise identified, are to pages of this article.

2. Only empirical statements can be confirmed.[1]

Since there must be a number of philosophers who, like Hempel, subscribe to both (1) and (2), it will be worthwhile to show that these theses are inconsistent—or at least that they are so except on an assumption which I think no one will be willing to make—and to consider ways of resolving their inconsistency.

Rather than try to show, however, that by supposing (1) we may derive the falsity of (2), I propose to proceed indirectly. I will show that Hempel's own definition of confirmation, at any rate, contradicts (2); and that it does so, just because it is a purely logical definition.

This requires that I should first state Hempel's definition, along with such explanations as its somewhat technical character calls for.

2

Confirmation is defined by Hempel for the statements, or, as he prefers to say, sentences, of a language which he refers to as ' L '. His characterisation of L, and its intended interpretation, we will return to shortly.

Hempel first defines a relation which he calls ' direct confirmation ', thus:

If M and S are sentences of L, then M *directly confirms* S if and only if (*a*) M is a molecule and (*b*) either the development of S for M is not analytic or S is analytic, and (*c*) M entails the development of S for M.

Then, partly in terms of direct confirmation, he defines confirmation, thus:

If M and S are sentences of L, then M *confirms* S if and only if: (*a*) M is a molecule; (*b*) there is a class K of sentences of L such that (i) K entails S and (ii) for every sentence T which is a member of K, M entails or directly confirms T.[2]

Two words in these definitions call for explanation.

First, ' molecule '. A molecule is any sentence of L which is ' either atomic or consists of atomic sentences and statement connectives'.[1] An atomic sentence of L consists of 'a predicate, followed by a parenthesised expression which consists of as many individual constants—separated by commas—as the degree of the predicate requires'.[2]

---

[1] Thesis (2) paraphrases part of the second sentence of the article, which opens (p. 122) as follows : ' The concept of confirmation occupies a central place in the methodology of empirical science. For it is the distinctive characteristic of an empirical hypothesis to be amenable, at least in principle, to a test based on suitable observations and experiments. . . .' Hempel's ' Studies in the Logic of Confirmation', *Mind*, **54**, 1945, opens with the same thought.

[2] See p. 142. In giving these definitions, I have translated Hempel's symbolism back into ordinary English ; and I have shortened his ' development ' idiom.

Second, 'development'. What is the development of a sentence 'for' a molecule—or, to give it its full name, the development of a sentence for the (finite) class of individual constants occurring (essentially) in a molecule? If the sentence contains no quantifier, this development is, for any molecule, the sentence itself. If the sentence does contain a quantifier, its development for a given molecule is the sentence that results when any universally quantified matrix the sentence contains is replaced by the conjunction of its substitution instances for all the individual constants occurring (essentially) in the molecule; and any existentially quantified matrix is replaced by the disjunction of its substitution instances for all the individual constants in the molecule.

Hempel defines confirmation, we saw, only for sentences of the language L. But the only restrictions he imposes on L are of a purely syntactical character, as follows.

First, its sentences must be formed, according to the rules of the lower functional calculus, out of signs for: statement connectives (for negation, disjunction, etc.); universal and existential quantifiers; (primitive) predicate constants of any finite degree; (primitive) individual constants and individual variables; parentheses and commas. Second, 'the rules of inference for L are those of the lower predicate calculus without identity sign'.[3]

It will be evident that, as Hempel remarks, 'many different languages satisfy [these] stipulations'[4]; and this fact will prove important. But the *intended* interpretation of L is expressly[5] that in which its predicates are interpreted as *observational* ones, and the individual constants are thought of as naming individuals of which observational predicates are predicable. In short, L is meant to be thought of as a language of *empirical* science, and Hempel's 'molecules', in particular, correspond to observation-reports.

In order to have before us an example of the kind of 'theorem' on confirmation which the definition yields, let S be the sentence '$(x)(Fx \supset Gx)$' and M be the sentence '$Fa . Ga$' 'F' and 'G' being logically independent predicates. Then under Hempel's definition of confirmation the theorem holds that M confirms S. For M is a molecule[1]; the development of S for M, viz. ' Fa $\supset$ Ga ', is not analytic; M entails ' Fa $\supset$ *Ga* '; hence M directly confirms S; whence the theorem.

---

[1] p. 125. [2] p. 125. [3] p. 124. [4] p. 123 [5] p. 126

It is this theorem which entitles us to say, for example, that under Hempel's definition ' This (material object) is a raven and is black ' confirms ' All ravens are black '.

It is important to realise, however, that this way of speaking, though not objectionable, is elliptical. For the theorems on confirmation that the definition yields can refer only to sentences of L. Now L, of course, is not really a language, only a language-schema.[2] It consists exclusively of ' sentences ' which, like ' Fa . Ga ' and ' (x)(Fx $\supset$ Gx) ', and unlike the sentences about ravens, are, obviously, not sentences but sentence-schemata. Hence, that ' Fa . Ga ' confirms ' (x)(Fx $\supset$ Gx) ' is a theorem under Hempel's definition; but that ' This (material object) is a raven and is black ' confirms ' All ravens are black ' is so only in an extended sense.

3

Hempel's definition of confirmation will be admitted to be a purely logical one. I now argue that it contradicts (2), just for that reason.

Thesis (2) will be false if it is possible for a statement which is not empirical to be confirmed. Now this must be possible under Hempel's definition of confirmation, unless it is a necessary consequence of that definition that if S is confirmed by some molecule then S is empirical. But it is obvious by inspection of the definition that there is nothing in it of which this could conceivably be a necessary consequence.

On the contrary, in the same sense, and for the same reasons, that it is a consequence of the definition that ' This (material object) is a raven and is black ' confirms ' All ravens are black ', it is also a consequence of it that ' This (number) is prime and is odd ' confirms ' All primes are odd '; and that ' This (immaterial object) is saved and is within the Church ' confirms ' There is no salvation outside the Church'. Thus arithmetical and metaphysical statements are confirmed under the definition. Similarly, moral generalisations, and logical ones, will be confirmed under the definition, given suitable choices of predicates and of things for the individual constants to name.

---

[1] Strictly, to comply with the definition of a molecule we should write M as ' F(a) . G(a) '.

[2] cf. p. 124

Thus, Hempel's definition allows non-empirical statements to be confirmed, that is, it contradicts (2).

Why does it have this result? The answer is obvious. The reason why the definition must fail to confine confirmation to empirical statements is that the definition is in purely logical terms, *whereas, whatever is a purely logical term, ' empirical ' is not.* (That a confirmed statement must be empirical cannot, of course, follow merely from the requirement that a confirmed statement be a sentence of L; for the empirical interpretation of L is just one of many possible interpretations.)

Consequently, any other purely logical definition of confirmation will likewise conflict with (2),[1] and theses (1) and (2) are inconsistent.

4

Hempel's definition, it so happens, also contradicts (2) for another reason, quite independent of its purely logical character, if we hold with Hempel[2] and nearly all other philosophers, that every analytic statement is entailed by any statement. For then it will follow from the definition that if S is analytic, S is confirmed (because directly confirmed) by M, for any molecule M.

5

Both these sources of conflict with (2) could be removed from Hempel's definition if we were to make certain amendments to that definition, and if we were prepared in addition to make a certain crucial assumption.

The definition already prevents self-contradictory sentences from being confirmed, except by self-contradictory molecules.[3] Hence, if we were to add to the definition the requirement that M be consistent, this would have the effect of confining confirmation to consistent sentences of L. If we were further to add to the definition the requirement that S be not analytic, the definition would then confine confirmation to sentences of L which are consistent and not analytic. And these requirements—that M be consistent, and that S be not analytic—are clearly of such a kind that the addition of them to the definition would not impair its purely logical character.

---

[1] For example, the identification, made by Professor Carnap and Professor Popper among others, of confirmation with ' favourable relevance ' in Keynes's sense. See the remark on Carnap in Section 6 below.

[2] That Hempel holds this is clear, for example, from the fact that he takes his adequacy-condition 3.3 to entail the condition 3.31. See pp. 127-128.

[3] Since it demonstrably satisfies Hempel's adequacy-condition 3.21.

Now the confinement of confirmed sentences to those which are consistent and not analytic, *if we were to assume that every statement, or significant sentence, other than a self-contradictory one, is either analytic or empirical*, would have the effect of confining confirmation to empirical statements.

Thus this assumption would be *sufficient* to render Hempel's definition (as amended) consistent with (2). But it is an assumption which will find few adherents, perhaps none, nowadays. I think it could never have found any, had philosophers realised that it embodies—what we saw in section 3 above would be *necessary* in order to reconcile a logical definition of confirmation with (2)—*a purely logical definition of empiricalness*. For surely this is a conception which needs only to be entertained in order to be dismissed.

I conclude, then, that theses (1) and (2) can be reconciled only on an assumption which we may safely exclude.

6.

The inconsistency we have discussed is not confined to Hempel's article on confirmation. Most writing on inductive inference in recent decades suffers from a less exact but easily recognisable variant of the same inconsistency.

For we find in this writing the assumption, corresponding to thesis (1), that inductive inferences are to be characterised and evaluated in terms of their form; but also the assumption, corresponding to thesis (2), that inductive inference is a topic peculiar to the philosophy of the *empirical* sciences. That these assumptions are inconsistent, I hope is evident from what has been said earlier.

Neither assumption is often explicitly made. But if the second were not made at all, it would be inexplicable why the discussion of inductive inference is confined to the books and chapters that it usually is. And if the first were not made, the ways in which induction is discussed would have to be very different from what they usually are. We could not, as we now do, discuss induction by discussing certain forms of, or certain rules for, inference. Nor could we discuss it, as is often done, by discussing stock examples: for in the absence of this assumption, nothing that was said about inferences concerning black ravens would have the least bearing on inferences concerning white swans.

Both of these inconsistent assumptions are at least suggested by the phrase 'inductive logic', as used for example by Professor Carnap in his great work on probability. For there he throughout refers to himself as engaged in constructing a system of inductive logic. Yet surely it is clear that he was constructing, rather, a system of probability-logic: a system which *becomes* an inductive logic, a logic of *empirical* science, when but only when his 'state-descriptions' are interpreted as possible objects of *empirical* knowledge. There is nothing in Carnap's characterisation of his 'languages' L_N, etc. (as there is nothing in Hempel's characterisation of L) to confine the predicates of L_N to observational predicates, or the individual constants to physical individuals; though no doubt that is the interpretation intended.

By way of contrast, it is instructive to recall that J. M. Keynes took for granted that in discussing inductive inference he was discussing at one and the same time the reasoning which might lead an ornithologist to conclude that all ravens are black, or Newton to the binomial theorem. In other words he took for granted the falsity of the second assumption now commonly made. And he did so precisely because he took for granted the truth of the first, that inference is inductive, or not, in virtue of its form.

7

I return now to the inconsistency of theses (1) and (2), to consider how best to resolve it.

The solution I propose is that (1) should be retained and (2) modified.

A modification of (2) which recommends itself on independent grounds, and which appears to render it consistent with (1), would be the thesis:

(2′) Only empirical statements can be confirmed *by observation-statements.*

I say that this modification recommends itself on independent grounds, because it accords well with the idea that confirmability-by-observation-statements is a criterion of the empiricalness (*not* of the *significance*) of non-analytic statements.

The other alternative would be to modify (1). We might, for example, define in purely logical terms a certain relation—we might even use Hempel's definition, but call the defined relation ' satisfaction ', say—and then proceed to identify confirmation with satisfaction when the satisfied statement meets certain non-logical requirements (such as the requirement of being empirical).

The fundamental objection to this course is that it requires us to abandon the idea which is behind thesis (1), and indeed behind most of what has been written about confirmation, induction, and logical probability: the idea that just as the conclusiveness of demonstrative inferences depends only on their form, so too does the partial conclusiveness, when they have any, of non-demonstrative inferences.

Philosophy Department
University of Sydney

ORIGIN AND CONCEPT OF RELATIVITY (III)*

G. H. KESWANI

1 *Relativity of all Motion*

IN parts I and II of this paper[1] we mentioned Poincaré's doubts about the validity of the principle of relativity *generally*. He found the principle valid for displacements in space and for uniform motion of translation but saw incontrovertible evidence for absolute effects in the case of accelerated and rotational motion. For example, even uniform rotational motion, say, of the earth, could be detected with Foucault's pendulum, by observations made on the earth itself, i.e. not relative to any other body. Poincaré reflected: Experiments have reference not to space but to bodies. Therefore, which are those bodies with respect to which bodies that turn, turn? Is it, he conjectured, that to apply the law of relativity in all its rigour, it must be applied to the entire universe? If we did so, how could the principle of relativity remain valid any more? If motion, in general, must be defined with reference to the universe of ' fixed ' stars, then it is surely absolute, and cannot be relative. Such were the questions that troubled Poincaré[2] in 1902.

Although Poincaré returned to the subject of relativity of motion again and again, the first serious attempt to extend the principle of relativity to all types of motion was made by Einstein,[3] culminating in his paper of 1916. Einstein thought that he had succeeded in this attempt through the ' principle of equivalence ', postulating equivalence of gravitational fields and accelerated co-ordinate systems, and the ' principle of covariance ' showing that it is possible to generalise the metric for *all* co-ordinate systems (CS) into a single invariant quadratic form, and to these ideas we now turn.

---

* Received 18.i.65

[1] This *Journal*, 1965, **15**, 286 and **16**, 19.

[2] *Science and Hypothesis*, Dover, pp. 77, 84, 113, 114, 169–172.

[3] *The Foundations of the General Theory of Relativity*, *The Principle of Relativity*, Dover, p. 111.

2 *The Principle of Equivalence*

The following CS are considered in various formulations of the principle of equivalence.

S_0, an inertial CS (we shall comment on inertial CS later).
S_1, a CS having acceleration (α) relative to S_0.
S_2, a CS falling freely in a gravitational field of *any* intensity.
S_3, a CS held at rest in a gravitational field (which can produce acceleration α).

The systems S_0 and S_2 are considered equivalent in regard to 'the motion of particles *or any other physical process*'[1] and likewise S_1 and S_3 are considered physically equivalent between themselves.[2]

In one of Einstein's formulations,[3] S_0 is considered equivalent to S_1 with a gravitational field (α) superimposed on it. The system S_1 is then regarded by Einstein as 'at rest', and equivalent to the inertial system S_0. We consider this formulation. Three questions now arise.

> (i) How is a gravitational field to be produced to change S_1 from a non-inertial to an inertial system so that S_1 may then be regarded as 'at rest', as Einstein proposed?

Obviously, a gravitational field can be brought into play only by introducing suitable masses. Can this be done without contradicting the relativity-thesis? Let us go back to the vexing rotating sphere of Poincaré. What is the system of bodies which will produce a gravitational field such that the rotating co-ordinate system can be regarded as 'at rest'? According to a calculation made by Thirring,[4] a mass in the shape of a concentric shell, rotating with the same angular velocity but in the direction opposite to that of the inner sphere under consideration, will induce the same centrifugal and Coriolos field in the inner sphere as the inner rotating sphere alone does, provided it has such enormous proportions that

$$MK/c^2R \simeq 1 \text{ or } M/R \simeq 1{\cdot}35 \times 10^{28} \text{ gm. cm.}^{-1}$$

---

[1] W. Pauli, *Theory of Relativity*, Pergamon, 1958, p. 145.

[2] R. C. Tolman, *Relativity Thermodynamics and Cosmology*, Oxford, 1949, pp. 174-175. Einstein, op. cit. p. 114.

[3] A. Einstein, *The Meaning of Relativity*, Methuen, 1950, p. 56.

[4] H. Thirring, 'Über die Wirkung rotierender ferner Massen in der Einsteinschen Gravitationstheorie', *Physik. Z.*, 1918, **19**, 33; 1921, **22**, 29. For a succinct remark on these papers, see: W. Pauli, op. cit. pp. 174-175.

where M is the mass of the outer concentric shell,
R is the radius of the shell,
K is the gravitational constant,
and c is the velocity of light *in vacuo.*

As we shall see, for a mass to have M/R of the order c^2/K, it has to be the universe itself. We might now be tempted to say that this only shows rather definitely the relativity of rotational motion; we may regard the sphere as rotating relative to the universe or the universe rotating relative to the sphere with the same angular velocity in the opposite direction. Either of them, when rotating, subject each particle of the inner sphere to the same field of acceleration. But this is not correct. The universe is one, while the rotating bodies are many and their axes of rotation and velocities varied. Are we to regard the universe as undergoing rotations of all kinds simultaneously, each schoolboy spinning his top, producing a fresh rotation? Also, how are we to rotate the masses around to convert a non-inertial rotating CS into an inertial one? The task appears to be operationally impossible. Moreover, rotation of the outer material of the universe will induce enormous centrifugal and Coriolos forces in it, which have not been observed. The correct way is to regard rotation as motion relative to the masses of the universe, producing as a consequence the observed centrifugal and Coriolos field. And motion relative to the universe and physical effects of rotation are both absolute.

> (ii) Is the equivalence between S_1 with a gravitational field superimposed, and the inertial CS, S_0, *complete*, as Einstein[1] asserted?

This is not so even according to what Einstein himself taught us. If a gravitational field exists, the Riemann-Christoffel (R-C) tensor cannot vanish at all points of the space-time, but for accelerated systems not subjected to a gravitational field, the R-C tensor vanishes at *all* points of the space-time, which is then flat. Equivalence cannot therefore be complete. Eddington[2] had clearly pointed out that the principle does not always hold good, only sometimes. A high degree of physical equivalence is, of course, to be expected since the equivalent gravitational field is so chosen as to produce the same kinematical effects as the acceleration field it replaces or nullifies.

---

[1] Einstein always regarded the equivalence to be complete, i.e. for all events. See, for example, *Ideas and Opinions* by Albert Einstein, Crown Pub., 1954, p. 287. In his original paper he said that a gravitational field could be 'produced' merely by changing the system of co-ordinates, i.e. by having suitably accelerated co-ordinate systems. *The Principle of Relativity*, Dover, p. 114.

[2] A. S. Eddington, *The Mathematical Theory of Relativity*, Cambridge, 1930, pp. 40-41. There is, however, a serious difficulty here. The trouble is that one does not always know whether or not the R-C tensor enters explicitly. We owe this remark

(iii) Even if there were complete equivalence between the two, would it establish relativity of all motion?

It is said Yes, because it is argued that by internal observations we cannot distinguish between a state of non-uniform motion and a gravitational field.[1] Quite apart from the consideration whether the curvature tensor does or does not figure in a given physical situation, further reflection will show that this assumed ambiguity in deciding whether it is mechanical acceleration or gravity, cannot lead to relativity of all motion. It appears to us that one can always distinguish between the two possibilities. A cosmonaut can know exactly whether it is an increase in the gravitational pull from some body or acceleration, and this by observations made in his capsule. He can measure the momentum of the escaping gases relative to the capsule and arrive at the acceleration of his capsule.

We may fittingly round off this discussion with a comment by one of the foremost living ' relativists ':[2]

> It appears as if general theory contained within itself the seeds of its own conceptual destruction because we can construct ' preferred ' coordinate systems. These preferred coordinate systems are not ' flat ' but they are determined by the intrinsic conditions of the physical situation. . . .

Our own point of view, as we explained earlier, lies farther in the same direction. Even ' flat ' space-time of accelerated CS, points towards an absolute frame of reference, and more about this in the sequel.

3 *Einstein's Theory of Gravitation*

Einstein's general theory of relativity is really only a theory of gravitation and of modifications in the laws of physics in gravitational

---

to one of the referees. A connected question is whether strictly *homogeneous* gravitational fields can exist. Also see: F. A. E. Pirani, *Lectures on General Relativity* (Brandeis Summer Institute in Theoretical Physics), Prentice Hall, 1965, p. 260.

[1] R. C. Tolman, op. cit. pp. 176–179. A. Einstein, *The Principle of Relativity*, Dover, p. 114.

[2] P. G. Bergmann, ' Observables in General Relativity ', *Rev. Mod. Phys.*, 1961, 33, 514.

fields. Even the principle of covariance, as Fock [1] has pointed out so emphatically, does not make it in any sense a general theory of relativity. Here we give an outline of the theory partly to see its connection with the notion of general relativity of motion and partly to get insight into the notion of an inertial CS. In this process, we shall develop a static cosmological model consistent with principal astronomical observations, purely from the field equations without using the cosmological term, and verifying Mach's principle exactly in a somewhat novel manner.

Einstein postulated an invariant 4-dimensional quadratic metric of space-time, a generalisation of Minkowski's metric of the special theory, as

$$ds^2 = g_{\mu\nu}(x)\, dx_\mu\, dx_\nu \quad (\mu,\nu = 1, 2, 3, 4) \tag{1}$$

where x's are the co-ordinates of the 4-dimensional manifold. It is assumed that form (1) has the signature $-3+1 = -2$. As Forsyth [2] has pointed out vigorously, this is a stupendous assumption, without mathematical justification. But it seems to 'work' physically. The 4 dimensions of 4-dimensional geometry are *co-ordinate* among themselves. Not so are the 3 dimensions of space and the 4th dimension of $\sqrt{-1} \times$ time.

The coefficients $g_{\mu\nu}(x)$, all functions of x's, depend on the gravitating masses present and their motions. Einstein regarded the space-time manifold as non-Euclidean,[3] $g_{\mu\nu}$ then defining the metrical properties of this manifold. How do the components of $g_{\mu\nu}$ vary with the distribution of matter? From $g_{\mu\nu}$, which constitute a covariant tensor, by differentiation of $g_{\mu\nu}$ only, we may build up a mixed tensor of the 4th rank (R-C or curvature tensor):

$$B^{\epsilon}_{\mu\nu\sigma} = \{\mu\sigma,\alpha\}\{\alpha\nu,\epsilon\} - \{\mu\nu,\alpha\}\{\alpha\sigma,\epsilon\} + \frac{\partial}{\partial x_\nu}\{\mu\sigma,\epsilon\} - \frac{\partial}{\partial x_\sigma}\{\mu\nu,\epsilon\},$$

where $$\{\mu\sigma,\alpha\} = \tfrac{1}{2} g^{\alpha\lambda}\left(\frac{\partial g_{\mu\lambda}}{\partial x_\sigma} + \frac{\partial g_{\sigma\lambda}}{\partial x_\mu} - \frac{\partial g_{\mu\sigma}}{\partial x_\lambda}\right), \text{ etc.}$$

---

[1] V. Fock, *The Theory of Space Time and Gravitation*, Pergamon, 1958, pp. xviii, 350, 370-374; 'The Principle of Relativity and of Equivalence in Einsteinian Gravitation Theory', *K. Norske Vidensk. Selsk. Forhandl* (Norway), 1963, **36**, 16.

[2] A. R. Forsyth, *Geometry of Four Dimensions*, Cambridge, 1930, Preface. Eddington, who showed great concern for 'first principles' has left the following coy remark: 'The 3 minus signs with 1 plus sign particularise the world in a way which we could scarcely have predicted from first principles.' op. cit. p. 25.

[3] N. Rosen suggested that space-time be regarded as Euclidean, with $g_{\mu\nu}$ defining the gravitational field but having no connection with the geometry of space-time, *Phys. Rev.*, 1940, **57**, 147. It appears to us that Rosen's point of view is the more satisfactory. See for example Appendix.

If the R–C tensor vanishes, then co-ordinate systems can be found such that $g_{\mu\nu}$ are constant throughout.

This part of the analysis is purely mathematical. Einstein then postulated that the law $B^{\epsilon}_{\mu\nu\sigma} = 0$ describes space-time without a gravitational field. The R–C tensor may be contracted ($\epsilon = \sigma$), giving the Ricci tensor:

$$G_{\mu\nu} = \{\mu\sigma,\alpha\}\,\{\alpha\nu,\sigma\} - \{\mu\nu,\alpha\}\,\{\alpha\sigma,\sigma\} + \frac{\partial}{\partial x_\nu}\{\mu\sigma,\sigma\} - \frac{\partial}{\partial x_\sigma}\{\mu\nu,\sigma\}.$$

Einstein *assumed* that $G_{\mu\nu} = 0$ is the law of gravitation in empty space surrounding matter. $g_{\mu\nu}$ are then given by the solutions of partial differential equations $G_{\mu\nu} = 0$. Since $g_{\mu\nu}$ and $G_{\mu\nu}$ are symmetrical tensors of 2nd rank, each of them has 10 independent components and it would appear that there are ten equations from $G_{\mu\nu} = 0$, to determine the ten unknown $g_{\mu\nu}$, but there are four identical relations between $G_{\mu\nu}$, leaving only 6 independent equations. In general, four additional 'co-ordinate conditions' are required. It appears that there is no way of telling which conditions to choose. This situation is regarded as unsatisfactory by some.[1]

If the components of a tensor vanish in one system of co-ordinates, they vanish in all systems, i.e. they are covariant with respect to arbitrary transformations of co-ordinates. Einstein[2] believed that covariance implies a 'general postulate of relativity', but as pointed out by Fock[3], this is by no means the case. The basic point is that covariance only ensures that the *form* of a law is the same for all CS. In the case of invariant form (1), for example, physical laws dependent on $g_{\mu\nu}$ vary from CS to CS, the different CS can therefore be distinguished from one another by observations made within each CS and a doctrine that relative motion alone is significant cannot be maintained for all types of motion.

In order to find $g_{\mu\nu}$ in non-empty space, Einstein introduced the energy-momentum tensor, $T_{\mu\nu}$ or $T^{\mu\nu}$ or T^{ν}_{μ}. For a perfect fluid the mixed form, which is the simplest, is

$$T^{\nu}_{\mu} = (\rho_0 + p/c^2)\,\frac{dx_\mu}{ds}\,\frac{dx_\nu}{ds} - \delta^{\nu}_{\mu}\,p/c^2 \left\{\delta^{\nu}_{\mu} = \begin{matrix} 1 \\ 0 \end{matrix}\ \text{if}\ \begin{matrix} \nu = \mu \\ \nu \neq \mu \end{matrix}\right\} \tag{2}$$

---

[1] V. Fock, op. cit. p. 194. N. Rosen, op. cit. p. 149.

[2] *The Principle of Relativity*, Dover, p. 117.

[3] V. Fock, op. cit., p. 350. Ricci and Levi-Civita had shown earlier in 1901 that from any assumed law one can always derive another law in covariant form which does not differ from the assumed law in any observable consequence: E. Whittaker, *A History of the Theories of Aether and Elecricity, The Modern Theories* (1900–1926), Thomas Nelson, 1953, p. 159.

where ρ_0 is the scalar proper density,
p is the scalar proper pressure or stress,
and $\frac{dx_\mu}{ds}$ is the the 4-velocity of the fluid.

Under static conditions, only the $\frac{dx_4}{ds}$ component (equal to 1) of the 4-velocity survives and then

$$T^\nu_\mu = \begin{vmatrix} p/c^2 & - & - & - \\ - & p/c^2 & - & - \\ - & - & p/c^2 & - \\ - & - & - & -\rho_0 \end{vmatrix}$$

For our present case we may regard the pressure as arising from the radial gravitational interactions of the ' particles ' constituting the fluid. The divergence of T^μ_ν always vanishes and it therefore satisfies the principle of conservation. Other energy tensors could be constructed but if the law of gravitation is to be expressed by second order differential equations, there is no alternative.

The divergence of G^ν_μ does not vanish identically but that of another tensor, $G^\nu_\mu - \frac{1}{2}\delta^\nu_\mu G$, does. Einstein then gave the following ' field equations ' for finding g^ν_μ in matter-occupied space:

$$-8\pi K T^\nu_\mu = G^\nu_\mu - \tfrac{1}{2}\delta^\nu_\mu G \qquad (3)$$

where $G = \delta^\mu_\nu \, G^\nu_\mu$ is a scalar called the Gaussian curvature and K is the gravitational constant.

Here we are interested in a particular solution of eq. (3). We consider a sphere of perfect fluid (this means that only radial interactions can be transmitted and that the fluid has constant proper density) and follow Synge.[1] We impose no limit on the size or mass of this sphere. Firstly, we set up Gaussian polar co-ordinates $(\rho, \theta, \phi, \tau)$,

where ρ is the geodesic distance from the centre of the spherical mass,
θ, ϕ are polar angles,
and τ is the proper time measured at the centre.

The metric form is then:

$$ds^2 = \eta^2 d\tau^2 - \frac{1}{c^2}\left[d\rho^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)\right] \qquad (4)$$

where η and r are functions of ρ only.

[1] J. L. Synge, *Relativity: The General Theory*, North-Holland Pub. Co., 1960, pp. 226-229, 276, 287-289. For a similar approach, see T. Levi-Civita, *The Absolute Differential Calculus*, Blackie, 1947, pp. 408-414.

Consider the 2-space for which ρ and τ have fixed values, and θ and ϕ are current co-ordinates. On account of spherical symmetry, all points of this 2-space are equivalent and it has a constant intrinsic Gaussian curvature $1/r^2$. The metric on it is then

$$ds^2 = -r^2\,(d\theta^2+\sin^2\theta d\phi^2).$$

The 2-space has, therefore, an invariant area, $4\pi r^2$.

Obviously ρ and η could be regarded as functions of r, when the metric form (3) may be expressed as

$$ds^2 = e^{\gamma}d\tau^2 - \frac{1}{c^2}[e^{\alpha}dr^2+r^2(d\theta^2+\sin^2\theta d\phi^2)],$$

e^{α} and e^{γ} being functions of r. We thus have what Synge calls 'curvature co-ordinates' (r, θ, ϕ, τ), the last one, as before, being the proper time measured at the *centre* of the sphere.

The field equations (3) contain p and ρ_0 on L.H.S. and e^{α} and e^{γ} on R.H.S. as unknowns. Other $g_{\mu\nu}$ are unity or zero. It turns out that to determine the four quantities p, ρ_0, e^{α} and e^{γ}, (3) yields three equations only. This is to be expected because p and ρ_0 could be related. We then get

$$e^{-\alpha} = 1-qr^2$$

$$e^{\gamma} = \{(3\sqrt{1-qR^2}-\sqrt{1-qr^2})/(3\sqrt{1-qR^2}-1)\}^2$$

$$\text{and } p/c^2 = \rho_0\{(\sqrt{1-qr^2}-\sqrt{1-qR^2})/(3\sqrt{1-qR^2}-\sqrt{1-qr^2})\} \tag{5a}$$

$$\text{with } \quad q = \tfrac{8}{3}\pi K\rho_0/c^2 = 2MK/c^2R^3 \tag{5}$$

$$R = \text{radius of the sphere}$$

$$\text{and } \quad M = \text{mass of the sphere.}$$

The metric inside the sphere is then

$$ds^2 = e^{\gamma}d\tau^2 - \frac{1}{c^2}\left[\frac{dr^2}{1-qr^2}+r^2(d\theta^2+\sin^2\theta d\phi^2)\right]. \tag{6}$$

Reverting to the radial geodesics, we find

$$d\rho^2 = \frac{dr^2}{1-qr^2} \quad \text{or} \quad \left(\frac{dr}{d\rho}\right)^2 = (1-qr^2). \tag{7}$$

The solution of this differential equation is $r = q^{-\frac{1}{2}}\sin q^{\frac{1}{2}}\rho$ which is plotted in Fig. 1. As ρ increases, r increases to the maximum value $R_u = q^{-\frac{1}{2}}$ and then diminishes. Beyond this value, one-to-one correspondence does not exist between the events $(\rho, \theta, \phi, \tau)$ and the co-ordinate tetrad (r, θ, ϕ, τ). Therefore only the portion OP of the graph is physically admissible.[1]

---

[1] For points of similarity: see L. O'Raifeartaigh, *Proc. Roy. Soc.* (A), 1958, 245, 202.

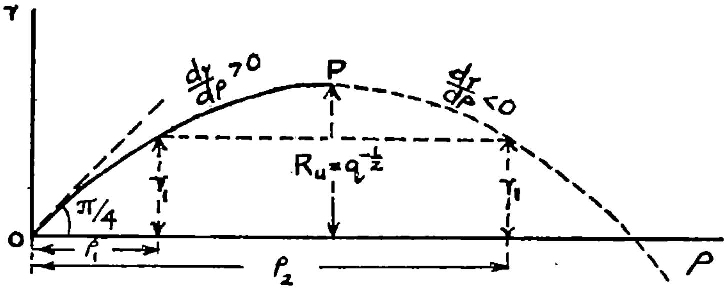


FIG. 1. Curvature coordinate/geodetic distance relationship: Same value of r (r_1) corresponds to more than one value of ρ (ρ_1, ρ_2, . . .).

4 *World Model*

We now identify R_u with the curvature of the whole 'universe'—of the biggest sphere (mass = M_u) accessible to the observer at the centre. The maximum geodesic distance observable is then

$$\rho_{\max} = \frac{\pi}{2}\, q^{-\frac{1}{2}} = \frac{\pi}{2}\, R_u = \frac{\pi}{2}\sqrt{\frac{3c^2}{8\pi K \rho_0}}$$

The coefficient $1/(1-qr^2)^{\frac{1}{2}}$ of dr becomes imaginary for $r>q^{-\frac{1}{2}}$. Therefore if $r>q^{-\frac{1}{2}}$, only the portion $R_u>r>0$ is physically accessible.

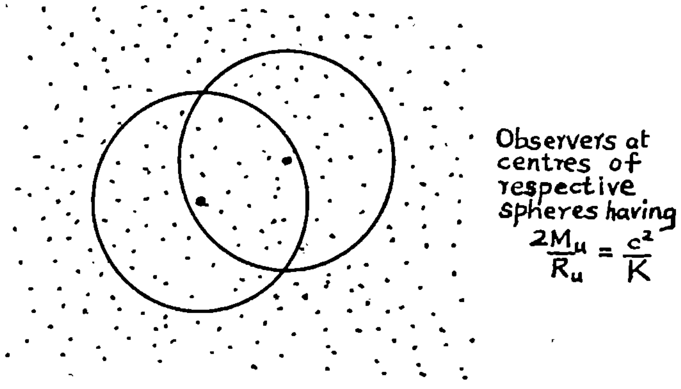

Observers at centres of respective spheres having
$\frac{2M_u}{R_u} = \frac{c^2}{K}$

FIG. 2. Universe open to observation.

However, for each observer, the 'universe' available for observation is *different*. The cosmology is then of the type indicated in Fig. 2. As we shall see presently, these ideas cohere physically.

The average density of matter in the universe[1] probably lies between

---

[1] J. A. Wheeler, *Lectures in Theoretical Physics*, ed. by W.E. Brittin *et al.*, Interscience Pub., 1963, p. 507.

$(1/3)\times 10^{-30}$ and 3×10^{-30} gm. cm.$^{-3}$. The energy density of neutrinos is not known and we take the upper value, 3×10^{-30}, and get

$$\rho_{\max} = 3{\cdot}8\times 10^{10} \quad \text{and} \quad R_u = 2{\cdot}4\times 10^{10} \text{ light-years.}$$

Putting $\frac{1}{q} = R_u^2$, eq. (6) becomes

$$ds^2 = \left(1-\frac{r^2}{R_u}\right)d\tau^2 - \frac{1}{c^2}\left[\left(1-\frac{r^2}{R_u^2}\right)^{-1} dr^2 + r^2(d\theta^2+\sin^2\theta d\phi^2)\right]. \qquad (8)$$

For a photon travelling radially from $r = R_u$ to the centre (observer), $ds = 0$ and

$$(1-r^2/R_u^2)\, d\tau^2 = dr^2/c^2\,(1-r^2/R_u^2)$$

and
$$\tau = \lim_{r\to R_u}\int_r^0 dr/c^2(1-r^2/R_u^2) = -\infty.$$

No photons from $r = R_u$ can therefore reach the observer. For $r<R_u$, the emitted and received wave-lengths are related as follows.

$$\frac{\lambda_{\text{emitted}}}{\lambda_{\text{received}}} = \frac{e^{\gamma/2}\text{ evaluated at emitter}}{e^{\gamma/2}\text{ evaluated at receiver}}$$

$$\therefore \lambda_{\text{received}} = \lambda_{\text{emitted}}\,(1-r^2/R_u^2)^{-\frac{1}{2}} = \lambda_{\text{emitted}}\left(1+\frac{r^2}{2R_u^2}+\frac{3}{8}\frac{r^4}{R_u^4}+\ldots\right)$$

and shift
$$\delta\lambda = +\lambda_{\text{emitted}}\times\frac{r^2}{2R_u^2} \quad \text{for} \quad \frac{r}{R_u}\ll 1.$$

This red-shift is obtained purely on account of the fact that light has to travel through a universe having the metric of (8) relative to the observer. The law of shifts is quadratic in the first approximation and not linear. Hawkins's[1] analysis of the data published by Humason, Mayall and Sandage shows that for field galaxies, the law $\delta\lambda/\lambda \propto r^{2\cdot 2}$ best fits the data for galaxies brighter than 14 magnitudes and that in general the results are more consistent with the quadratic law than with the linear.

We have been quietly using a metric of the same form as de Sitter's, but this metric has now been obtained in an entirely natural way from the field equations, as a result simply of recognition of a singularity[2] of these equations. We must now discuss a difficulty of the de Sitter universe which arises in the present case as well. 5 (a) for the 'pressure' or radial interaction reduces to

---

[1] G. S. Hawkins, 'Expansion of the Universe', *Nature*, 1962, **194**, 563–564.

[2] This singularity is, of course, the same as the famous Schwarzschild singularity. For the exterior metric also, the singularity arises at $2m/R = 1$ or $2MK/c^2R = 1$, or $M/R\sim 10^{28}$ gm/cm. For no known object (fundamental particle, quasar, galaxy) is $2m/R>1$. The object would be invisible—but detectable in certain cases—if $2m/R>1$. It is an interesting conjecture whether, for certain particles (for example, neutrinos) $2m/R>1$. Rosen and Fock get an isotropic and unique metric-form in which the singular surface lies at $m/R = 1$.

$$p/c^2 = -\rho_0. \tag{9}$$

The pressure is constant but negative and equal in magnitude to $c^2 \times$ density. It had long been believed that the notion of such a negative pressure was inadmissible. The way has now been cleared for the admissibility of (9) by the work of Jordan,[1] McCrea,[2] McVittie,[3] Sciama[4] and Davidson.[5] But the development here has, of course, nothing to do with the creation theories of cosmology. What (9) says is that 'the gravitational energy of any particle arising from its interaction with the rest of the universe is exactly equal and *opposite* to the energy due to the mass itself'. On this account, to use McVittie's[6] words, we have a 'gravitationally steady state'. As A. Hass and P. Jordan suggest, we may put the equation $M_u K/c^2 R_u = 1$, where M_u is the mass of the universe, in the form $KM_u^2/R_u = M\ c^2$, which in Jordan's[7] words means, 'that the negative potential energy of gravitation for the whole universe is equal to the sum of the rest energies of the stars'. It may appear that a negative pressure may, somehow, offend Newtonian mechanics. McCrea[8] has pointed out that if the pressure is *uniform*, its gradient is zero, and it will, therefore, neither appear in the equation of motion nor in the equation of continuity. This stress cannot therefore be observed.

The relationship $p = -\rho_0 c^2$, or 'gravitational energy of a mass equals minus its inertial energy', may then be looked upon as a neoteric verification of Mach's principle.

5. *Inertial CS*

An inertial CS is usually defined as a system in which 'force-free' bodies are not accelerated with reference to the CS. However,

---

[1] P. Jordan, 'Formation of the Stars and Development of the Universe', *Nature*, 1949, **164**, 637–640.

[2] W. H. McCrea, 'Relativity Theory and Creation of Matter', *Proc. Roy. Soc.* (A), 1951, **206**, 562–575.

[3] G. C. McVittie, 'A Model Universe Admitting Interchangeability of Stress and Mass', *Proc. Roy. Soc.* (A), 1952, **211**, 295–301.

[4] D. W. Sciama, 'On the Origin of Inertia', *M.N.R.A.S.*, 1953, **113**, 34.

[5] W. Davidson, 'General Relativity and Mach's principle', *M.N.R.A.S.*, 1957, **117**, 212–224.

[6] G. C. McVittie, op. cit., Summary, p. 295.

[7] P. Jordan, op. cit. We do not discuss here a discrepancy of a factor equal to 2.

[8] W. H. McCrea, op. cit. p. 571.

Einstein[1] himself pointed out that this definition involves a circular argument: 'A mass moves without acceleration if it is sufficiently far removed from other bodies; we know that it is sufficiently far from ther bodies only by the fact that it moves without acceleration.'

Obviously, with Einstein's theory of gravitation before us we cannot ignore the masses of the universe. However (8) shows that at the observer ($r = 0$) the metric is

$$ds^2 = d\tau^2 - \frac{1}{c^2}[dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)],$$

which is locally flat with $g_{\mu\nu} = \text{diag}\ [1, -1, -1, -1]$. Deviation from flatness due to sun's gravity at Pluto is found to be 10^{11} times greater than that due to the masses of the universe. Therefore CS at rest with reference to the universe are force-free and inertial.[2] Do other inertial systems in a state of motion relative to the system of the stars exist? It is an observed fact that systems in uniform motion of translation relative to the universe are also inertial, if the definition is the one given at the beginning of this section, but this does not necessarily imply that CS in uniform motion relative to the universe are all equivalent for the description of the laws of physics. Although force-free material bodies may not be accelerated relative to such CS, there may be other effects of uniform translation and this we shall examine in the sequel. That brings us to the special theory of relativity again.

6 *Special Theory and Lorentz Transformation Equations*

As we saw in Part II of this paper, Einstein used: (1) the principle of relativity, and (ii) the principle of constancy of velocity of light, to establish Lorentz transformation equations (L.T.E.).

We now propose to show (not for the first time!)[3] that L.T.E. can be derived from the second postulate alone, without recourse to the principle of relativity. The novelties and sometimes weird results predicted by L.T.E.—and verified experimentally—really arise from the second postulate and are independent of the notion of relativity. L.T.E., no doubt, leave certain quantities invariant, for example the wave equation of electromagnetic propagation. Experiments dependent on electromagnetic propagation will, therefore, naturally conform to the relativity-principle, but there is no ground, really, for asserting that L.T.E. will conform to the relativity hypothesis *always*.

---

[1] *The Meaning of Relativity*, Methuen, 1950, p. 57.

[2] The following dichotomy mentioned by R. H. Dicke is thus resolved for CS at rest w.r.t. the universe. '... on the one hand we consider inertial effects to be related in an intimate way to the rest of the universe, and on the other hand we assume that the universe has no observable effects on the laboratory and on the numerical content of physical laws.' 'Eötvös Experiment and the Gravitational Red Shift', *Amer. J. Phys.* 1960, **28**, 345.

[3] For a similar demonstration, see: G. Stephenson and C. W. Kilmister, *Special Relativity for Physicists*, Longmans, 1958, pp. 9–10.

Let S (x, y, z, t) and S$'(x', y', z', t')$ be two inertial frames as defined earlier, with S$'$ moving with velocity v relative to S in the direction of X-X$'$ axes which coincide. Let Y$'$ and Z$'$ axes be respectively parallel to Y and Z axes. The problem is to find the transformation equations, i.e. x', y', z' and t' as functions of x, y, z, t and v. We expect these equations to be linear, such as $x' = g_{10}+g_{11}x+g_{12}y+g_{13}z+g_{14}t$, where g_{ik} are some coefficients involving v. This simply means that the physical content of transformations is independent of the origin of the co-ordinates. The total differential $dx' = g_{11}dx + \ldots$ is the same throughout, which would not always be the case with non-linear transformations. By a suitable choice of the origins, g_{10} may be made zero. Y$'$ and Z$'$ axes are similarly located w.r.t. the X$_0$ axis and the direction of v. Obviously x' and t' are independent of y and z, and y' and z' independent of x and t. By shifting the origin of S$'$, Y and Y$'$ axes may be made to coincide. Therefore y' is independent of z. Likewise z' is independent of y. These judgements are based on the cylindrical symmetry of the problem and Euclidean geometry. Finally, therefore, these considerations lead to the following transformations

$$\left.\begin{array}{l} x' = g_{11}x+g_{14}t, \quad y' = g_{22}y, \quad z' = g_{22}z, \\ t' = g_{41}x+g_{44}t. \end{array}\right\} \tag{10}$$

g_{22} is a scale factor independent of v. We now make use of the second postulate. Since the velocity of light is independent of the velocity of an observer relative to the source of light, the spherical wave-front of a light pulse leaving the coincident origins of S and S$'$ at time $t = 0 = t'$ will have the equations in the two systems as given in (11) and (12) below.

$$x^2+y^2+z^2-c^2t^2 = 0 \tag{11}$$

$$x'^2+y'^2+z'^2-c^2t'^2 = 0 \tag{12}$$

Since $y' = g_{22}y$ and $z' = g_{22}z$, the conditions that (12) should transform into (11) are the same as the ones for $(x'^2-c^2t'^2)/g_{22}$ to transform into $(x^2-c^2t^2)$. We consider first, transformation of $(x'^2-c^2t'^2)$ into $(x^2-c^2t^2)$. Substituting for x' and t' from (10), we have

$$(g_{11}x+g_{14}t)^2-c^2(g_{41}x+g_{44}t)^2 = x^2-c^2t^2.$$

This is an identity in the variables x^2, xt and t^2, and it furnishes three equations for the four unknowns g_{11}, g_{14}, g_{41} and g_{44}. Leaving aside some tedious algebra, the solutions are found to be:

$$\begin{aligned} x' &= g_{22}(x-\psi t)/(1-\psi^2/c^2)^{\frac{1}{2}} \\ t' &= g_{22}(t-\psi x/c^2)/(1-\psi^2/c^2)^{\frac{1}{2}}. \end{aligned}$$

The coefficient g_{22} has been restored. ψ is some function of v, which is the physical parameter characterising the problem. We may put ψ as a power series, thus:

$$\psi = a_0+a_1v+a_2v^2+\dots$$

Now for $v = 0$, $x' = x$.

$$\therefore\ x' = g_{22}(x-a_0t)/(1-a_0^2/c^2)^{\frac{1}{2}} = x, \text{ for all values of } t.$$

Obviously then,

$$g_{22} = 1 \quad \text{and} \quad a_0 = 0.$$

The transformation equations thus, simplify into the form:

$$\begin{aligned} x' &= (x-\psi t)/(1-\psi^2/c^2)^{\frac{1}{2}}, \quad y' = y, \quad z' = z, \\ t' &= (t-\psi x/c^2)/(1-\psi^2/c^2)^{\frac{1}{2}}, \\ \psi &= a_1v+a_2v^2+\dots \end{aligned}$$

These equations will transform a spherical pulse of velocity c in S′ into a spherical pulse of the same velocity in S. Further simplifications in the above equations do not appear possible without further assumptions. We now suppose that these equations are valid for all events. We can hardly expect different equations for electromagnetic and other phenomena. This is *not* tantamount to the assumption of relativity or equivalence of all inertial CS. We might recall that in pre-relativity days we expected evidence counter to relativity, if Galilean transformations held good for electromagnetic phenomena. In fact, at the present stage, we may nurse the expectation that application of the new transformations to mechanics may now show that the principle of relativity is not valid for mechanics.

Consider now the events (0, 0, 0, t'), i.e. the time recorded by a clock (mechanical or electromagnetic) at the origin of S′. For this clock,

$$x' = 0 = (x-\psi t)/(1-\psi^2/c^2)^2 \quad \text{or} \quad \psi = x/t.$$

Now x/t for this clock is its relative velocity v as measured in S. The transformation equations are then found to be the usual ones:

$$\begin{aligned} x' &= \beta(x-vt), \quad y' = y, \quad z' = z, \\ t' &= \beta(t-vx/c^2) \quad \text{with} \quad \beta = (1-v^2/c^2)^{-\frac{1}{2}}. \end{aligned} \tag{13}$$

7 *Second Postulate*

The second postulate was questioned many years ago by Ritz.[1] He suggested—what has recently been revived by Dingle—that light is emitted with constant velocity relative to the source or emitter 'just as though it were a material particle fired by an unvarying mechanism', as Dingle [2] puts it.

Firstly, we must clearly recognise that if the second postulate goes, L.T.E. go with it too. Lorentz-transformation is now so intimately woven into the warp and woof of modern physics that if it was invalid, this would surely have been detected somewhere.

And, what do we exactly mean by 'constant velocity relative to the emitter'? It is rather difficult to say what emits light. Nucleons may have velocities $\sim 0\cdot 1c$ and orbital electrons $\sim 0\cdot 01c$. In the case of light, it is not at all easy to define 'velocity relative to the emitter'.

It appears to us that Dingle is driven to reject the second postulate because he feels that although there is overwhelming evidence for relativity, L.T.E. contradict it. The fact of the matter is that we have evidence for L.T.E. and only to that extent for relativity. No more. Indeed Einstein [3] himself later took the view that 'the *whole content* of the special theory of relativity is included in the postulate that the laws of nature are invariant with respect to a Lorentz transformation'.

It is then obvious that if a 'law of physics' could be discovered which was not invariant under a Lorentz transformation, the principle of relativity would be violated. Here we must note a distinction. If the space and time measures of inertial CS are generally different but constant within various systems (as they are with L.T.E.), by experiments carried out within the respective systems, differences in the laws of physics are not to be expected. Basically then, two types of relativity may be mentioned. The 'pure' variety—to use Sherwin's[1] word for it—in which all inertial CS are identical, not only in regard to the laws of physics as observed within any CS but also when a *comparison* of the laws in different CS is made. And 'apparent' relativity, in which although the laws of physics are seemingly the same within various CS, a suitable comparison of two or more systems may reveal differences in these laws. The question then is, how this comparison may be made. A comparison of lengths cannot, perhaps, be made meaningfully. Could the time-measures be compared? Let us see what is to be expected theoretically from L.T.E. and what the experimental results are.

---

[1] W. Ritz, *Collected Works*, pp. 317, 427, 447. Also see W. Pauli, op. cit. pp. 5–9.

[2] H. Dingle, 'On Inertial Reference Frames', *Science Progress*, 1962, **50**, 5. Dingle has taken the view that there is no reliable direct experimental proof of the second postulate. In support, J. G. Fox has argued that in the older observations purporting to support this postulate, light from moving sources passed through intervening media (air, glass, envelope of gas) and not vacuum, and that according to the extinction theorem of light, these media became the secondary sources of emission, etc. W. R. Haseltine has pointed out that extinction of light in the intervening media does not occur if the signals from the primary source have a velocity different from c. More recent experiments avoiding Fox's objection, have confirmed the second postulate (D. Sadeh, T. Alvager *et al.*, J. F. James and R. S. Sternberg) excepting an experiment by W. Kantor. P. Burcer has shown that the 'relativistic' theory of Kantor's experiment is in conformity with his result.

[3] *Ideas and Opinions* by Albert Einstein, 1954, Crown Pub., p. 370. Also see p. 283.

8 *L.T.E. and Time of Different CS*

Consider two contiguous events associated with a point moving with velocity v relative to S. Let their measures in S be $[x, t]$ and $[x+\delta x, t+\delta t]$. Transformed by L.T.E. the corresponding events in S$'$ are

$$[x', t'] = [\beta(x-vt), \beta(t-vx/c^2)],$$

and

$$[x'+\delta x', t'+\delta t'] = [\beta(x+\delta x-vt-v\delta t), \beta(t+\delta t-vx/c^2-v\delta x/c^2)].$$

We note that (x, t) and (x', t') are the measures within the respective systems.

$$\therefore \delta t'/\delta t = \beta(1-v^2/c^2) \quad \text{or} \quad \delta t'/\delta t = \beta^{-1}, \quad (\delta x/\delta t = v).$$

We consider the limit as δt and $\delta t' \to 0$, i.e., the case when the two events coincide in the limit. Then:

$$dt'/dt = \beta^{-1}.$$

This is a familiar result, but what does it mean? Why is the result $dt'/dt = \beta^{-1}$ and not vice-versa? The explanation is that in deriving L.T.E. we have assumed one system (S) to be at rest relative to which the other system moves with velocity v. Otherwise we should not have got this asymmetrical relationship. We may remark that the above result follows in a straightforward manner from L.T.E. but how can it be regarded as in conformity with relativity, or the notion of equivalence of the two CS? t and t' are the time measures within two

---

[1] C. W. Sherwin, 'Some Recent Experimental Tests of the "Clock Paradox"', *Phys. Rev.*, 1960, **120**, 18. He remarks, 'Thus whatever its theoretical difficulties may be, pure relativism is untenable experimentally'.

CS. Is the law of passage of time not a law of physics? It is, but this law is not invariant under a Lorentz transformation. It is to be noted, however, that the time measure of all devices is equally affected in S′, so that the form of laws within this system will remain the same. Einstein himself asserted later that the time-measures of different inertial systems are different and that moving clocks change their rhythm. Writing jointly with Infeld[1] he said, 'In classical mechanics it was tacitly assumed that a moving clock does not change its rhythm. . . . We can well imagine that a moving clock changes its rhythm, so long as the law of this change is the same for all inertial CS. . . . We must accept the concept of relative time *in* every CS. . . . Every CS must be equipped with its own clocks at rest, since motions change the rhythm. . . .' We should like to emphasise that it could not be the 'rhythm' or 'time' as it 'appears' or is measured from the 'other' system. A change in rhythm of this kind (first order Doppler shift in frequency) had long been known and caused no trouble in understanding. Also, no meaning can be attached to the description 'time of S′ as measured in S'. All difficulties have arisen because relativity has wrongly been identified with L.T.E.

Let us clear up some irrelevant subtleties. There is nothing in our story about clocks as such. We talked only of the time-measures, t and t', of two CS. *No observations from one system on another are involved.* One need not fuss about the 'simultaneity of distant events'. We consider the limiting case when two events coincide and thus get a relationship at the point of intersection of two world lines. Again, one might protest: S′ could as well have been chosen as at 'rest', leading to the perfectly reciprocal relationship $dt/dt' = \beta^{-1}$. One could have, but obviously it is nonsense to say $dt'/dt = \beta^{-1}$ and $dt/dt' = \beta^{-1}$ $(\beta \neq 1)$. One makes a last-ditch stand and says that only the relative velocity enters into L.T.E. This is true because no other velocity was considered, but we can proceed consistently only if it is conceded that v is measured relative to an absolute system.

The question then is: which is the physical system that is always at rest? The system of the fixed stars qualifies for this. The time measures (t') of all other inertial systems are related to the time measure (t) of the system of the fixed stars (S) by the equation $dt' = dt/\beta$.

If we have two systems S′ and S″, so that S″ has velocity v relative to S′ and S′ has velocity w relative to the system of the universe S, the time-dilation factors for the two systems are as follows (for simplicity we assume w to be in the same direction as v; this is the worst case for the present argument, the result is easily generalised):

---

[1] A. Einstein and L. Infeld, *The Evolution of Physics*, Cambridge (Paperback), 1961, pp. 184, 186, 189.

$$[1-w^2/c^2]^{-\frac{1}{2}} \text{ for S}',$$

and
$$\left[1-\left(\frac{v+w}{1+vw/c^2}\right)\Big/c^2\right]^{-\frac{1}{2}} \text{ for S}''.$$

The relative time-dilation factor between S′ and S″ is then

$$\left[\frac{1-v^2/c^2}{(1+vw/c^2)^2}\right]^{-\frac{1}{2}}, \tag{14}$$

the time-processes being slower in S″. Now, in various experiments carried out on the earth, w, the velocity of the earth relative to the universe, is certainly less than $10^{-3}c$ as indicated by the present degree of accuracy in the observation of red-shift of distant nebulae. If w was higher, a definite anisotropy would appear with preponderance of blue-shifted galaxies in one direction.[1]

It appears to us that no experiment has yet been carried out to test (14) for its denominator. As we remarked earlier, experiments dependent on electromagnetic propagation, of necessity, fully comply with the relativity-condition. Also experiments like that of Tomlinson and Essen [2] cannot discriminate between the 'pure' and 'apparent' relativistic points of view. The vibrating rod in their experiment essentially functioned as a clock which was contained in one CS (that of the earth). Orientation of a clock in the same CS has, of course, no effect on its rhythm. In the experiments on μ mesons, $v/c \simeq 1$, and the accuracy is still about 9 % only.[3] It must be noted that slowing of time is observed to take place in the fast-moving system of the mesons and *not* in the slower system of the earth having much smaller velocity relative to the universe. It might be said that in certain recent experiments using the Mössbauer effect and involving velocities v of the order $10^{-6}c$, absolute velocity of the earth, if any, should have shown up. We shall presently see that these particular experiments could not, in theory, have made this possible.

---

[1] R. H. Dicke, op. cit. p. 344.

[2] G. A. Tomlinson and L. Essen, *Proc. Roy. Soc.* (A), 1937, 158, 606–633.

[3] D. H. Frisch and J. H. Smith, 'Measurement of the Relativistic Time Dilation Using μ Mesons', *Amer. J. Phys.*, 1963, 31, 342.

9 *Experiments Using Mössbauer Effect*

Two kinds of such experiments pertinent to the present inquiry have been performed. In one, performed by Hay *et al.*,[1] and repeated with greater accuracy by Kündig,[2] a change in the characteristic frequency of absorption of gamma rays by an absorber situated near the outer periphery of a rotor was observed, the source of gamma rays being near the centre of the rotor. The gamma rays passing through the absorber reached a stationary counter outside the spinning rotor. The counting rate was observed to increase with the angular velocity of the rotor. In the other experiment, first performed by Pound and Rebka,[3] the rate of absorption of gamma rays by an absorber was found to change with the temperature of the emitter, on account of changes in the characteristic frequency of emitting nuclei participating in the thermal velocities of the crystal lattice.

Synge[4] has given a beautiful demonstration to show that for a source at radius R′ and the absorber at radius R *anywhere* on the rotor, the ratio of the frequency of the source, ν', and of the absorber, ν, is

$$\frac{\nu'}{\nu} = \left(\frac{1-R^2\omega^2}{1-R'^2\omega^2}\right)^{\frac{1}{2}}$$

exactly, ω being the angular velocity of the rotor. The shift to be expected is dependent on the magnitude of the radii R′ and R and of the angular velocity but *not* on the *relative* velocities of the source and the absorber. Thus, when the source and absorber lie at the two ends of a diameter, the expected shift is zero, although the relative velocity of the two is $2\omega R$. Indeed in an experiment performed earlier, this theoretical result was verified![5]

In the second experiment also, the changes in temperature result in changes in the accelerations of the emitting nuclei in harmonic motion. (In the theory of the experiment worked out by Josephson,[6] it is assumed that the forces coupling the atoms are harmonic.) In fact the situation is similar to the rotor experiment except that the acceleration ($-\omega^2 R' \sin \omega t$) of an emitting nucleus is variable here but is constant ($-\omega^2 R'$) in the rotor experiment.

---

[1] H. J. Hay *et al.*, 'Measurement of the Red-Shift of an Accelerated System Using the Mössbauer Effect in Fe^{57}', *Phys. Rev. Letters*, 1960, 4, 165.

[2] W. Kündig, 'Measurement of Transverse Doppler Effect in an Accelerated System', *Phys. Rev.* 1963, 129, 2371.

[3] R. V. Pound and G. A. Rebka, 'Variation with Temperature of Energy of Recoil-free Gamma Rays from Solids', *Phys. Rev. Letters*, 1960, 4, 274.

[4] J. L. Synge, 'Group Motions in Space-time and Doppler Effects', *Nature*, 1963, 198, 679.

[5] D. [illegible] Champeney and P. B. [illegible] Ray Source and Detector on Sam[illegible].

[illegible]

Concluding Remark

The principle of constancy of velocity of light leads to Lorentz transformation equations (L.T.E.), without the use of the principle of relativity, the two being independent schemes. An analysis of L.T.E. shows that a high degree of equivalence in the description of the laws of physics exists for different inertial systems. The principle of relativity, therefore, assumes verisimilitude by identification with L.T.E. Further analysis, however, shows that although the time measures of different inertial systems are constant within, they vary from system to system. This forces us to recognize the system assumed as at rest in deriving L.T.E. to be a primary preferential system. The system of the masses of the universe is the obvious choice for this. In fact, the absolute effects observable in the case of accelerated motion demand this primary system.

Room No. 28A, 4th Floor,
Baroda House, Annexe I
Curzon Road
New Delhi, India.

APPENDIX

Motion of Mercury's Perihelion

The present argument is based on the consideration that the 3-space is non-Euclidean according to Einstein, $g_{\mu\nu} dx_\mu dx_\nu$ defining the metric of space (*The Principle of Relativity*, Dover, p. 161; *The Meaning of Relativity*, Methuen, p. 59), while with Rosen, $g_{\mu\nu}$ merely define the physical state, without any connection with the geometry of space which is regarded as flat. When the prevailing geometry is non-Euclidean, the ratio circumference/radius is not 2π and in the present case it is in fact less than 2π.

The equation of the orbit is obtained from the exterior Schwarzschild [illegible]

Various symbols have the meaning given in section 3 of the present paper and m has the meaning given below. The exact orbit is given by an elliptic function but, as usual, we take the following trigonometric equation of the orbit which is sufficiently exact.

$$1/r = [1+e\cos(\phi-\tilde{\omega}-3m^2\phi/h^2)]m/h^2 \qquad (2)$$

where $m = \mathrm{M.G}/c^2$ = mass of the sun $\times\, \mathrm{G}/c^2$,

$h = r^2 d\phi/ds$ (which turns up as a constant of integration),

e = eccentricity of the orbit,

and $\tilde{\omega}$ = longitude of the perihelion.

The period of successive perihelia is then $2\pi/(1-3m^2/h^2) \simeq 2\pi(1+3m^2/h^2)$. If one revolution corresponded to 2π this result would mean that the perihelion of Mercury advances by $6\pi m^2/h^2$ per revolution of Mercury or by 43 sec. per terrestrial century.

However, the quantity 2π arising above is a number and may not correspond to one revolution in a non-Euclidean space. Indeed R. C. Tolman (*Relativity Thermodynamics and Cosmology*, Oxford, 1949, p. 208, footnote) and G. C. McVittie (*General Relativity and Cosmology*, Chapman & Hall, 1956, pp. 87-88), both consider the fact of non-Euclidean geometry but conclude that no significant error is introduced by ignoring this fact. They show that whether one takes (i) r or $\int \frac{dr}{(1-2m/r)^{\frac{1}{2}}}$, (ii) s the proper time at Mercury, or τ the co-ordinate time at the origin (or for that matter the proper time measured by an astronomer on the earth) in the calculation, the end result is not sensibly affected. But the fly in the ointment is ϕ.

We now calculate the ratio circumference/radius. In fact we only carry McVittie's calculation forward. (A similar method is indicated by L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, Pergamon, 1959, p. 307). Using the interior and exterior Schwarzschild metrics, we have the radial geodetic distance

$$\rho = \int_0^{\mathrm{R}_\odot} (1-2mr^2/\mathrm{R}_\odot^3)^{-\frac{1}{2}} dr + \int_{\mathrm{R}_\odot}^{\mathrm{R}} (1-2m/r)^{-\frac{1}{2}} dr \qquad (3)$$

where, $\mathrm{R}_\odot$ = radius of curvature at the surface of the sun (taken as bounded),

and R = radius of curvature at the mean position of Mercury with reference to the origin (centre of the sun).

$\mathrm{R}_\odot$ and R may, without sacrifice of the necessary accuracy, be identified with the astronomer's measures of the radius of the sun and the average distance of Mercury from the centre of the sun, respectively. Assumption of the average distance, again, does not affect the result materially. Expanding (3) we get

$$\rho = \int_{0}^{R_\odot} (1 + mr^2/R_\odot^3 + \ldots)dr + \int_{R_\odot}^{R} (1 + m/r + \ldots)dr \simeq R + m(\tfrac{1}{3} + \log_e R/R_\odot)$$

Then, the angle swept out per revolution is

$$2\pi R/[R + m(\tfrac{1}{3} + \log_e R/R_\odot)] \simeq 2\pi - 7{\cdot}65 \times 10^{-7}.$$

The second term represents departure from Euclidicity. And in one terrestrial century, the deficit of $7{\cdot}65 \times 10^{-7}$ per revolution, amounts to 65 sec. Taking care of the signs, the centennial advance is found to be $43 + 65 = 108$ sec.

According to Rosen, r is the radial distance itself and $(1-2m/r)^{-\frac{1}{2}}$, etc., would be the components of the gravitational field which enter into the dynamics of motion but do not alter the space-geometry. Thus the eq. (2) has the usual interpretation leading to the result of 43 sec. In fact Rosen finds the metric to have the unique form

$$ds^2 = \frac{1 - m/r}{1 + m/r} d\tau^2 - \frac{1}{c}\left[\frac{1 + m/r}{1 - m/r} dr^2 + (r+m)^2 (d\theta^2 + \sin^2\theta d\phi^2)\right].$$

One possible criticism of this demonstration may be answered. Why take the radial geodetic distance ρ in the calculation? Because, the increment of radial distance is $d\rho$, the radial metric being

$$ds^2 = (1 - 2m/r)d\tau^2 - \frac{d\rho^2}{c^2}, \text{ etc.}$$

Regarding appropriateness of Gaussian co-ordinates in the present case, also see: *Introduction to General Relativity*, R. Adler *et al.*, McGraw Hill, 1965, pp. 59–62.

ERRATA

'Origin and Concept of Relativity', Parts I and II

Part I (**15**, No. 60, 1965)

p. 286, line 15: For 'is' *read* 'are'

p. 288, line 5: For 'references 3 and 4' *read* 'references 2 and 3, p. 287'

p. 304, line 21: For '$u_x\rho' = \beta(u_x - v)$' *read* '$u'_x\rho' = \beta(u_x - v)\rho$'

Part II (**16**, No. 61, 1965)

p. 26, line 21: For 'word-line' *read* 'world-line'

Read *k* for K on pp. 31 and 32, except in the figure on p. 31, and line 7 on p. 32, where *read* *K* for K. Also, *read* *k* for k in the figure on p. 31.

PHYSICS AND COMMON SENSE*

NICHOLAS MAXWELL

1 *Physicalism and Common Sense*

THE basic problem to be discussed is this: To what extent, in what sense, may the discoveries of physics legitimately conflict with our ordinary common sense views about the world based on our ordinary experience? In order to discuss this question, I consider a particular extreme view about the nature of the world—a view I call Physicalism —which may seem to be supported by the findings of physics, and which asserts at least that:

(*a*) The world is made up entirely of only a few different sorts of things—the fundamental physical entities.

(*b*) Precise, exceptionless laws govern the way in which these entities change.

(*c*) Human perception is almost entirely deceptive: almost all perceptual qualities, e.g. qualities such as colours, sounds and smells, have no real, no objective existence.

According to physicalism, only those qualities which apply to the fundamental entities, or to aggregates of fundamental entities, such as mass, position, electric charge, etc., really exist.

Many eminent physicists appear to have believed in some form of physicalism. Thus Planck has written that certain '. . . considerations . . . and not any logical argument . . . compel us to assume the existence of another world of reality behind the world of the senses; a world which has existence independent of man, and which can only be perceived directly through the medium of the world of the senses, and by means of certain symbols which our senses allow us to apprehend. It is as though we were compelled to contemplate a certain object in which we are interested through spectacles of whose optical properties we were entirely ignorant.'[1] Again, Einstein has written: 'Physics is an attempt conceptually to grasp reality as it is thought independently

* Received 15.vii.65

[1] M. Planck, *The Universe in the Light of Modern Physics*, Allen & Unwin, 1931, pp. 8–9

of its being observed. In this sense one speaks of "physical reality"'[1], and: 'The belief in an external world independent of the perceiving subject is the basis of all natural science. Since however, sense perception only gives information of this external world or of "physical reality" indirectly, we can only grasp the latter by speculative means.'[2]

The findings of physics do not of course oblige us to adopt some version of physicalism. We may argue that scientific investigation involves an extension, elaboration and refinement of common sense, of our ordinary experience, and therefore that scientific discoveries cannot contradict in any fundamental way the tenets of common sense that are based on ordinary experience. According to this second view—which I call the 'common sense' theory—scientific discoveries can only undermine those of our ordinary views about the world that are based on inadequate or distorted observation.

Thus scientists may find some means to extend the range of our experience, to that which is very distant perhaps, or very small, or to that which occurred long ago. As a result of obtaining this new, direct or indirect, observational evidence, the scientist is in a position to correct previous assumptions made about the nature of that part of the universe in question.

Again, it is the physicist's fundamental aim to discover, in so far as this is possible, those very general regularities between phenomena to which there are no exceptions; to discover, in other words, the 'laws of nature'. On the whole physicists have been extraordinarily successful in this search. But from our immediate, uninformed experience it is not at all obvious that the nature of the world is such as to be amenable to this kind of investigation. Thus from ordinary human experience, not altogether unreasonable conclusions may be reached about the nature of things in the world, which may seem to be rendered highly implausible by the subsequent apparent discovery of universal laws of nature—by the discovery that it is possible to formulate laws which (*a*) are not refuted by experience (perhaps within certain limits) and (*b*) make accurate predictions possible.

But none of this implies that there is anything inadequate or distorted about our observation of the familiar 'furniture of the earth'. Thus, according to this second view, as long as we are not mad, drugged, blind, deaf or dreaming, our ordinary experience provides us with knowledge about the visual, auditory, tactile, etc., properties of the things on this earth which no legitimate scientific theory could ever refute. Roses really are red, dogs really do growl, etc.

---

[1] P. Schilpp (ed.), *Albert Einstein: Philosopher-Scientist*, Tudor, New York, 1957, p. 81

[2] A. Einstein, *The World as I see It*, Bodley Head, 1935, pp. 156-157

These two theories, the ' common sense ' theory and physicalism, have of course been formulated with no very great precision. Any number of different versions of each theory may be developed. Enough has been said however, I hope, to make quite clear the following crucial point: we have here two very different theories about the nature of the world, for either of which the claim may be made that it is *scientific*, that it is supported by science, and takes into account possible future scientific discoveries.

I wish to stress that these are theories about the nature of the world, theories about *things*, and not primarily theories about the nature of science or the theories of physics (although each theory may, incidentally as it were, imply different things about the nature of science). Thus, according to physicalism, the world is made up solely of a few different kinds of fundamental physical entities, nothing really being coloured, etc.; according to the ' common sense ' theory, the world is at least made up of all the different kinds of things that we ordinarily experience, many of which really are coloured, etc.

We have here surely an extraordinary state of affairs. Physics appears to be a particularly precise discipline; most physicists at any one time seem to agree as to which theories are acceptable, which unacceptable. Yet we have just seen that wholly divergent views may be held about what physics can tell us about the nature of the world.

The reason for this, briefly, is that precisely what we consider physics can tell us about the nature of the world depends to a considerable extent on the kind of interpretation we think can legitimately be given to the mathematical formalism of a physical theory. And on this point there is no general agreement.

Let us suppose a wholly successful, all-embracing, fundamental physical theory has been found which suits ideally the physicalist's purposes. We may suppose that this ' ideal ' theory is such that given any isolated system, and given the so-called ' initial conditions ' corresponding to some state of that system, then the theory will in principle enable us to predict all subsequent states of the system. Further we may suppose that the theory at least ostensibly postulates the existence of just a few different kinds of fundamental entities.

The physicalist would interpret such a physical theory in the following manner. Given any isolated system, in order to describe what really exists at any instant we require (*a*) the fundamental theory, which states in general terms what kind of things the fundamental entities are, (*b*) the initial conditions, which specify the precise values at the moment in question of such variable attributes of the fundamental entities as position, momentum. According to the physicalist, a description of what really exists at a given moment is only *complete* if it enables us to deduce descriptions of subsequent states of the system. By hypothesis, the above ' ideal ' theory does provide such a ' complete ' description.

The physical theory might however be given a quite different, non-physicalist interpretation. Briefly, it might be argued that the theory would not *really* postulate the existence of fundamental entities. The theory would *really* only describe a *model*, a conceptual tool invented by the physicist solely in order to facilitate the prediction of observations. Thus the fact a successful physical theory apparently postulates and describes fundamental physical entities in no way implies that such entities *really* exist.[1]

Defending such different interpretations to one and the same mathematical formalism clearly amounts to defending different theories about the nature of the world. The question arises: Does a successful mathematical formalism, given a physicalist interpretation, constitute a possible theory of physics, as opposed perhaps to a theory of metaphysics?

I accept here without argument Popper's solution to the problem of demarcation between physics and metaphysics: a theory, in order to belong to physics, must at least be experimentally falsifiable.[2]

From this requirement it follows that the kind of interpretation of the mathematical formalism of a physical theory, demanded by physicalism, which we may call ' tentative realism ', is legitimate. In practice of course no isolated system can be observed; but experiments can nonetheless be devised to test deductive consequences of the theories intended to apply to isolated systems.

It should be noted that even if one day the kind of ideal fundamental physical theory described above is found, the physicalist will still never know with certainty that entities precisely like those described by the theory do really exist, since it will never be possible to know with certainty that the theory is true. Nonetheless the theory *might* be completely true, and hence entities precisely like those postulated by the theory *might* really exist. The fact that the theory must be open to experimental refutation ensures that it is meaningful to call the theory false, which in turn ensures surely that it must be at least *meaningful* to call the theory *true*.

---

[1] Variants of this view have been defended by Berkeley, Mach, Poincaré, Bridgman, Heisenberg, Bohr, and by the logical positivists. For an exposition and criticism of this view, under the heading ' Instrumentalism ', see K. Popper, ' Three Views concerning Human Knowledge ', in *Conjectures and Refutations*, Routledge and Kegan Paul, London, 1963.

[2] K. Popper, *The Logic of Scientific Discovery*, Hutchinson, London, 1959

In giving a tentative realist interpretation to physical theories we are not obliged to assert that entities postulated by *all* acceptable physical theories really exist. Suppose a theory, by postulating certain entities, is able to make accurate predictions within certain conditions or limits, but outside these, goes wildly astray, and is, in other words, refuted. Clearly, in this case, not all of the entities postulated by the theory really exist. Nonetheless the theory might be retained, perhaps because it is the only theory available, or perhaps because of its relative simplicity.

The kind of tentative realist interpretation of physical theories indicated here is often opposed by a view which combines a non-realist interpretation of theories that apply to the so-called micro-level, with a realist interpretation of theories that apply to the macro-level. Such a view is prompted partly by the belief that scientific theories must be observationally verifiable (the kind of belief that is presupposed by Hume's polemic against causation as 'necessary connections' between events), partly by the fact that such a view enables one to side-step the apparent wave/particle paradoxes of micro-phenomena. Unfortunately there is insufficient space for a discussion of such matters here.[1]

The view that physicalism might conceivably be formulatable as a physical theory raises the problem: Why would not such a theory be refuted immediately by our ordinary experiences? Is not the theory that the world is in reality colourless, soundless, odourless, etc., refuted by the fact that we do see colours, hear sounds, smell smells?

I wish to argue that physicalism is only a defensible theory if interpreted in such a way that it does *not* imply that there are no colours, etc., in the world. The problem here is to interpret physicalism in such a way that it does not contradict an acceptable version of the 'common sense' theory. But before indicating how the 'common sense' theory and physicalism may be reconciled, I wish to discuss two versions of physicalism that reject outright the 'common sense' theory.

---

[1] For a criticism of the principle of complementarity (and thus of phenomenalist interpretation of quantum mechanics) see: M. Bunge, 'Strife about Complementarity', this *Journal* 6, 21; for a suggestion of how apparent wave/particle paradoxes of micro-phenomena may be resolved see A. Landé, 'Why do Quantum Theorists ignore the Quantum Theory?' this *Journal*, 15, 60.

The thoroughgoing physicalist clearly requires a theory of perception which explains why our ordinary experience does not refute physicalism. From Popper's demarcation requirement, it follows that if physicalism is to be formulatable as a physical theory, then from the theory, and a description of the constitution of the human brain and sense organs, it must be possible to deduce a theory of perception which:

1. Explains why most of our experience is delusive.
2. Indicates precisely what aspect of our experience is *not* delusive, so that predictions of the physicalist theory may be tested experimentally.

2 *Two Physicalist Theories Rejected*

The two physicalist theories of perception I wish to discuss are (*a*) dualism, (*b*) the discriminatory response/brain process theory, defended recently by Professor Smart.[1]

For the purposes of this discussion I shall call those qualities which, according to physicalism do really, objectively exist, *physicalist*, and those perceptual qualities which, according to physicalism, do not really, objectively exist, *phenomenal*. This leaves open the possibility that some, but clearly not all, physicalist qualities, such as for example shape, may also be perceptual.

(*a*) *Dualism*. This doctrine accounts for the alleged deceptiveness of human perception as follows. We know that if I perceive, let us say, a red rose, then a causal sequence of events takes place between the surface of the rose and my brain via my eyes and optic nerve. According to the dualist, it follows from this that my visual sensation of the rose must be an entity caused by, or in some way associated with, some event in my brain, but nonetheless distinct from any event in my brain, and open only to my own inspection. Hence, in perceiving the rose, I really only perceive my visual sensation of the rose. I can therefore have no reason to believe that the perceptual quality, redness, applies to the physical object I have called a rose: redness is a quality that can apply only to visual sensations that occur in the mind. Of course in practice if a physical object causes us to have the visual sensation of redness, then we will tend to call that physical object 'red'; but strictly it must always be *false* to ascribe a perceptual quality to a physical object. In general, physicalist qualities apply only to physical objects, perceptual qualities only to sensations or sense impressions, which exist exclusively in our minds. Dualism thus involves a serious addition to physicalism, since according to dualism the world is made up of the two entirely different sorts of things: (i) fundamental physical entities, (ii) sense impressions, or, more generally, ideas. Ideas exist only in dimensionless ' bubbles ', called ' minds ', which are associated in some strange way with certain clusters of fundamental physical entities called ' brains '.

---

[1] J. J. C. Smart, *Philosophy and Scientific Realism*, Routledge & Kegan Paul, London, 1963, Chaps. 4, 5

Such a dualist theory of perception does provide an explanation of how human perception might be deceptive, of how things might not really be as they seem. It is a mistake, however, to suppose that dualism can be derived from the existence of the causal sequence of events involved in any perception. We may stipulate that a necessary condition for perceiving an object is that a certain kind of causal sequence of events should take place between the object and the brain via the eyes and optic nerve. The visual sensation (as opposed to the perception) we may define as the last event in the above sequence, leaving open the question of whether this is a brain process or some peculiarly ' mental ' occurrence. From such an explication of ' perception ' it follows that far from *only* being able to perceive our own visual sensations, in practice we *never* perceive our visual sensations.

A more serious objection to dualism as a physicalist theory of perception is however that it does not meet the second of the above two requirements for such a theory. According to dualism, we can have no reason to believe that any experience is *non*-delusive, if our experience is interpreted as providing information about the physical world. This is because, according to dualism, we can *only* perceive our own sensations [illegible]e, it being impossible to perceive any physical object.[1] In other words, the assumption that dualism is true implies that there can be no evidence in favour of the theory. (More precisely, dualism implies that no singular existential proposition about the physical world can be verified, and hence that no physicalist theory can be experimentally refutable.)

(*b*) *The Discriminatory Response/Brain Process* theory of Professor Smart. Briefly, Smart's thesis is that phenomenal qualities must be understood not as intrinsic, unanalysable, objective properties of things, but in terms of the discriminatory responses of human beings. Colour, for example, is to be analysed in the following manner. First we explicate the phrase ‘ discriminate with respect to colour ’ without introducing colour as an undefined primitive. A person discriminates between objects with respect to colour if the discrimination can be made when the objects are illuminated with daylight but cannot be made when the objects are illuminated with monochromatic light. A normal human percipient with respect to colour is then defined as a person who is able to make all the colour discriminations that anyone else can can make. We then define objects of the same colour as objects which a normal human percipient would not be able to discriminate between with respect to colour. Thus Smart declares: ‘ We might say at a first shot, that ‘ this is red ’ means roughly that a normal human percipient would not easily pick this thing out of a heap of geranium petals, though he would easily pick it out of a heap of lettuce leaves.’[1]

---

[1] This, basically, is the point behind Berkeley's polemic against Lockean substance, see ' Three Dialogues ' in G. Berkeley, *A New Theory of Vision and other writings*, Everyman, London, 1910.

Smart acknowledges that this account of colour has one slight flaw: we can imagine that colours of all objects undergo a systematic change, so that for example lettuce leaves become red, geranium petals green. According to the above account, this radical change would not be detectable, since human beings would continue to make colour discriminations just as before. In order to meet this difficulty, Smart admits that we must take into account the inner experiences that accompany discriminatory responses. But this does not involve admitting that unanalysable mental or psychic entities exist, since these inner experiences are, according to Smart, nothing more than complicated neurophysiological events.

This last point—which we may call, following Smart, the ‘ brain process theory of inner experiences ’—does not, I think, represent a weak point in Smart’s position. The theory is based on the argument that in having an inner experience I am presented with no evidence for the existence of entities over and above brain processes, in the sense required by dualism. Thus, if I experience, for example, the visual sensation of a red patch, then I have undeniable evidence only that something exists which, in some unknown respect, resembles that which exists when I perceive a red patch. Hence all available evidence supports the hypothesis that the ‘ something ’ in question is a particular brain process. This does not imply however that ‘ brain process ’ and ‘ inner experience ’ have the same meaning.

---

[1] J. J. C. Smart, op. cit. p. 79

The brain process theory is often considered as an implausible last step in the physicalist's reductive programme. I shall assume here however, without further discussion, that this theory, if formulated and defended entirely independently of physicalism, is an entirely sensible, unobjectionable view, with considerable ordinary empirical evidence in its favour. I assume that it is in full accordance with our ordinary experience, and by no means implies that men are merely mechanisms, or that life can have no significance or value. In other words, I assume that the brain process theory, formulated independently of physicalism, is in full accordance with the ' common sense ' theory. It is, I think, in part just because the brain process theory is thought to presuppose physicalism that it is usually considered to be so implausible.[1]

Objections to Smart's physicalism lie elsewhere, in connection with his ' discriminatory response/brain process ' analysis of qualities such as colours. (It should be noted that acceptance of a version of the brain process theory does not automatically implicate us in an acceptance of the role which this theory plays in Smart's analysis of colours.) Crudely, the obvious objection to Smart's physicalism is surely this: the theory just does not account for the existence of qualities like *redness*. Whenever we see a red object surely we just are indubitably aware of the existence, somewhere in the world, of a unique, unanalysable quality which we may call ' redness '. Yet according to Smart it is just this kind of quality that does not exist at all, either in things or in our minds.

Unfortunately, in order to give a precise formulation to this apparently simple objection, it seems it is necessary to raise the following somewhat intricate considerations.

Smart does not always make it quite clear which of the two following positions he is defending:

(i) The physicalist analysis of sentences such as ' This is red ' explicates satisfactorily what we ordinarily mean by this sentence.

(ii) ' This is red ', as ordinarily understood, is always *false*; hence all perceptions of colours are delusive, and colours are only phenomenal qualities.

Clearly Smart must be defending one or other of these positions, but not both. It would seem that the following objection to the first position is decisive: an analysis of ' This is red ' in terms of discriminatory responses of certain biological mechanisms misses the most important part of what is ordinarily meant by this sentence, namely that unanalysable part which can only be understood by actually seeing a red object, or, at the very least, by experiencing the visual sensation of redness. To this Smart might reply ' But this experience can in turn be analysed in terms of some brain process, which clearly could be fully described without mentioning any such quality as " redness " '. The answer to this (accepting the brain process theory as formulated here) is that in order fully to understand ' This is red ' it is necessary to have the brain process in question occur in one's own brain; it is wholly irrelevant whether or not one is in possession of a description of this brain process. In having this brain process occur in one's own brain for the first time something wholly new is learnt; one discovers what red things are like, just that which the congenitally blind never know.

---

[1] For a defence of these points see N. Maxwell, *Physics and Common Sense, a critique of Physicalism*, M.A. thesis, 1965, Manchester University, pp. 72–126.

To this the following qualification must be added. It is just possible, as far as our knowledge goes at present, that the *public* meaning of ' This is red ' (i.e. that meaning which is common to all the rather different meanings which people *may* give to ' This is red ') can be completely analysed in Smartian terms. This would be the case if, from person to person, the inner experience that accompanied the perception of a red object was different. Suppose the brain process that accompanies my perception of a red object is A, while the brain process that accompanies your perception of a red object is B; suppose further that if B is induced by some means in my brain I experience a visual sensation which I regard as quite different from ' experiencing the visual sensation of redness '. In this case I would say ' What I mean by " This is red " is quite different from what you mean '. Nonetheless a part of what I meant by ' This is red ' would be the same as what you meant by this sentence. This ' public ' part would be explicated entirely by a Smartian analysis.

However, given these circumstances (which may very well not be the case) the *private* meaning which I give to ' This is red ' would still be a perfectly genuine meaning. For another person to understand this private meaning it would be necessary only for that person to have a brain process which I judged to be sufficiently like A in the relevant respect in order to constitute for that person ' experiencing the visual sensation of redness '. Thus arguments such as Wittgenstein's against the possibility of a private language,[1] do not apply here, since a definite procedure exists for determining whether two people have similar inner experiences.

To sum up, we must reject the contention that Smart's analysis of colours explicates all that we ordinarily mean by such sentences as ' This is red '.

---

[1] L. Wittgenstein, *Philosophical Investigations*, Blackwell, Oxford, 1958, §258, etc.

But no doubt Smart does not wish to make the above claim for his analysis of colours. Smart would maintain I think that a sentence such as ' This is red ', if given its ordinary meaning, is always false. In other words Smart is defending a physicalist theory of perception, according to which all perceptions of colours (as ordinarily understood) are false, i.e. according to which colours (as ordinarily understood) are phenomenal qualities only.

The question arises: Does this theory satisfactorily explain why all perceptions of colours are delusive, why colours are phenomenal qualities only? It is important to realise that Smart's discriminatory response/brain process theory does not in itself explain why certain perceptual qualities are only phenomenal. After all, to any perceptual quality whatsoever there will correspond certain discriminatory responses and brain processes of human beings. Hence merely from the existence of such discriminatory responses and brain processes we cannot deduce that the quality in question is only phenomenal, and does not really, or objectively, exist. Thus the theory is no more than a device whereby the term for a quality which has been shown to be phenomenal can be given a physicalist interpretation.

Why then does Smart maintain that colours as ordinarily understood are phenomenal? Basically Smart distinguishes physicalist from phenomenal qualities as follows: Physicalist qualities are just those which physicists will ascribe to the fundamental physical entities and to all possible assemblies of fundamental entities, once (or if) they formulate a satisfactory, basic, all-embracing physical theory. Phenomenal qualities are then those qualities which we do ordinarily attribute to objects and appear to perceive, but whose existence cannot be predicted from [illegible] a full knowledge of the properties of the fundamental entities. But the mere fact that a physical theory does not predict the existence of a perceptual quality cannot in itself be a sufficient reason for concluding that that quality does not really exist. In fact perception of a quality not predicted by a physical theory would *refute* that theory unless it could be shown that either (1) our perception of the quality is always delusive, or (2) the theory does not imply that the quality does not exist. Hence nothing that has been mentioned so far supports the hypothesis that all perceptions of colours are delusive.

Smart does provide certain *ad hoc* arguments in support of the hypothesis that qualities such as colours do not really, objectively exist, and therefore that our perceptions of colours are always delusive. For example he argues that the view that colours have an objective existence is rendered extremely implausible by the fact that the physical property which corresponds to any colour is usually complex and, from the point of view of physics, wholly arbitrary.[1] Thus, to light of any specific colour, there corresponds infinitely many mixtures of light of different wavelengths and intensities. But these considerations show only that from the point of view of physics, classification in terms of colours seems arbitrary; it does not follow that colours do not really, objectively exist. From the point of view of the person who perceives colours, classification in terms of colours is very far from arbitrary.

It is true that doubt may be raised as to what is to count as a veridical experience of colour, particularly by the findings reported by E. H. Land.[2] But any such doubt must presuppose in the end that *some* experience of colour is veridical.

3 *An Acceptable Physicalism*

I wish to turn now to an examination of our original problem: How can physicalism be interpreted so as to render it consistent with an acceptable version of the ' common sense ' theory?

In essence the solution I offer to this problem is very simple. Both theories provide comprehensive descriptions of the world. These two descriptions are however compatible, since they are not both the same *kind* of description. Thus each description applies to almost all there is, and yet does not tell us all that there is to know about that to which it applies. Our task in what follows will be to:

(*a*) Explicate the precise requirements which each description must fulfil.

(*b*) Show that an ideal, fundamental physical theory would support physicalism given this interpretation.

(*c*) Show that the two kinds of description are compatible.

(*d*) Refute such claims as that only physicalism describes the world as it *really*, *essentially* or *objectively* is.

(*a*) In saying what exists at any moment we can only say what that which exists is *like*, in some respect or other. A number of very different *kinds* of resemblances can be found between things. The physicalist and the ' common sense ' descriptions classify things in terms of different *kinds* of resemblances between things.

---

[1] J. J. C. Smart, op. cit. pp. 69–72

[2] *Scientific American*, May 1959, 200 (5), 84–99.

I suggest that physicalism is to be interpreted as providing the following kind of classification of things:

(i) That which exists at any instant is classified, in the simplest possible way (i.e. with the smallest possible basic vocabulary), in terms of causal consequences, i.e. is described in such a say that descriptions of what exists subsequently can be deduced. (In the terminology of physics, given an isolated system, that which exists at any instant is described in such a way that descriptions of subsequent states of the system can be deduced.)

(ii) Things are classified only in terms of those resemblances which any intelligent being, however its sensory equipment may be constructed, can discern, discover, become aware of.

It is assumed that these two requirements are compatible.

It should perhaps be emphasised that the first requirement makes the following presupposition: 'It is at least possible, or conceivable, that from a true description of what exists at one instant, true descriptions of what exists subsequently can be deduced.' (It is precisely this which is presupposed by a tentative realist interpretation of the kind of 'ideal' fundamental physical theory described on p. 297.) Hume of course rejected the above proposition. Hence in accepting the first requirement for a physicalist description (and, incidentally, in accepting a realist interpretation to physical theories), we are committed to denying the validity of Hume's defence of the proposition 'It is not possible, not conceivable, that from a true description of what exists at one instant, true descriptions of what exists subsequently can be deduced'.[1]

The ways in which 'common sense' descriptions classify things are less easy to specify precisely. In general, things are classified solely in terms of resemblances easily discernible to human beings, in terms of resemblances associated with the experiences, interests, and purposes of human beings. A 'common sense' description of an object may classify that object in terms of (amongst other things):

(i) What the object looks like, feels like, etc., to a human being.

(ii) The object's past (e.g. how the object was made, where it came from).

(iii) What the object is used for.

(iv) The object's causal properties.

Often these, and other, kinds of classifications are combined in any one description, although of course one kind may predominate.

---

[1] For a refutation of Hume on this point see N. Maxwell, op. cit. pp. 57-69

(*b*) It is clear enough, from our previous discussion, that a certain kind of legitimate fundamental physical theory would, if given a tentative realist interpretation, support physicalism in the above sense. Whether or not contemporary physics supports such a version of physicalism is, as I have already indicated, a controversial issue.

(*c*) It is further quite clear that it is possible for physicalism and the 'common sense' theory, given the above interpretations, to be compatible. If physicalism satisfies the above two requirements for a physicalist description, then qualities that are, for example, discernible only to beings with sensory equipment similar to those of human beings will fall wholly outside the province of the theory. Physicalism will neither imply that such qualities exist, nor imply that such qualities do not exist. At the most the theory would imply, given a physicalist description of a human being, that that human being will make certain kinds of discriminations in certain situations, and will have certain brain processes. Since physicalism does not imply the non-existence of the quality I discern in seeing a tomato say and call 'redness', it must be compatible with physicalism to say of an object 'This is red' in *this* sense.

It should be noted that the requirements for the physicalist description do not necessarily imply an altogether sharp distinction between physicalist and non-physicalist qualities. Thus it might be maintained that even such apparently typical non-physicalist qualities as colours are physicalist qualities, since they are (i) not altogether causally inefficacious, (ii) perhaps perceivable by any intelligent being if we permit sufficiently drastic brain surgery.

(*d*) We turn now to a refutation of the claims that only the physicalist description is *really* true, that only physicalism describes the world as it *really*, *essentially* or *objectively* is.

The question 'Is the world *really* as described by the "common sense" theory or as described by physicalism?' only makes sense if the two descriptions are incompatible, if only one can be true. But as we have already indicated, a comparison of the requirements for each description makes it quite clear that it is perfectly possible for both descriptions to be true. Hence the above question makes no more sense than: 'Is this rod *really* one foot or twelve inches long?' Both colours and electrons *really* exist, although they are very different sorts of things.

It may be argued that fundamental particles, and all aggregates of such particles, are *essentially* colourless, *essentially* without perceptual properties, precisely the reverse being true of anything to which the 'common sense' description refers. Hence the two kinds of description are incompatible: only one can be true.

But what is meant by the phrase 'the *essential* properties of a thing'? The following two related definitions may be given. The essential properties of an electron, for example, are (*a*) those which we consider to be the most important, or (*b*) those which a thing must possess if it is to be called an electron. Clearly in giving the essential properties of the electron, according to the second definition, we are giving no more than an explanation of what we mean by the term 'electron'. There will be a tendency to define 'electron' in terms of the properties which we consider to be the most important. For this reason the two definitions of essential properties will tend to be equivalent.

From this explication of the notion of 'essence', it follows that we may grant that the essential natures of electrons and stones say are different without thereby being committed to maintaining that the two kinds of descriptions are incompatible.

It is true that from a complete physicalist description alone it would be impossible to deduce the *perceptual* qualities of things, but this is due, not to the fact that things do not really possess perceptual qualities, but to the fact that the physicalist description is *incomplete*: it does not tell us all that there is to know about the world. It does not tell us what it is like to be a human being alive and experiencing in the world. In particular, it tells us little about that aspect of objects which we discover in direct perception.

Misunderstanding of this point may to some extent be responsible for the view, defended for example by R. Harré,[1] that stones and electrons belong to different ontological classes, that both stones and electrons exist, but in different senses of 'exist'. Certainly the 'essential' properties of stones and electrons are very different. Further, the epistemological status of these properties is in each case very different. Propositions that attribute perceptual properties to stones may be verified in a straightforward way by observation: propositions which attribute physicalist properties to fundamental physical entities are never verified; at most they survive all attempts to refute them. None of this implies however that stones and electrons exist in different senses of 'exist'. Once we have made quite clear what sort of thing a stone or an electron is, then it can only be either true or false that stones exist, either true or false that electrons exist.

---

[1] R. Harré, *Theories and Things*, Sheed and Ward, London, 1961, p. 85

Again, it may be argued that the physicalist kind of description is truer than other kinds because it is more objective. Thus Smart argues '. . . our ordinary manner of talking about the world is suffused with concepts which relate the things in the world to our human concerns and interests, and which depend, in often unnoticed ways, on our human physiology and our particular station in space-time. . . . If the anthropocentricity inherent in these concepts is not brought out into the open we can have a misleading picture of the world. We think, for example, of objective colour *qualia* or of an objective *now*, much as the Hebrews looked up at the dome of the sky and thought that this was a solid half-spherical shell, or firmament, and did not realise that the apparently solid object was an illusion of their own perspective.'[1]

We may agree with Smart that ' Our ordinary manner of talking about the world is suffused with concepts which relate the things in the world to our human concerns and interests, and which depend . . . on our human physiology '. As much is clear from the conditions for the ' common sense ' description. Whether or not it follows that such a ' manner of talking about the world ' is subjective cannot be decided until the distinction between subjectivity and objectivity has been defined (see below). But whatever our decision on this point, from the fact that the ' common sense ' description is ' suffused with concepts which relate the things in the world to our human concerns and interests, and which depend . . . on our human physiology ', it does *not* follow that such a description is in any way false, or that the corresponding qualities do not exist. Descriptions such as ' That is—a car, a Picasso, green, a waste-paper-basket, etc.', relate the things described to human concerns and interests, and can only fully be understood and verified by human beings, but are not, on that account, false. The fact that you must possess a human physiology in order to perceive the greenness of things does not imply that grass is not really green.

We may grant that if typically ' common sense ' descriptions are employed as if they meet both requirements for the physicalist description, then all such descriptions will be false. If colours, for example, are thought of as physicalist qualities, then it would be false to attribute colours to things. But we are not obliged to give such an interpretation to the concept of colour. We are not obliged to interpret the word ' red ' so that ' This is red ' is bound to be false.

Finally, a few words about the distinction between objectivity and subjectivity. I wish to suggest that the traditional manner in which this distinction is drawn makes an implicit appeal to Cartesian dualism, and must be rejected if we reject Cartesian dualism. It might be stated thus: a quality is objective if it exists independently of the observer, subjective if its existence depends on the existence of the observer. This has a clear enough meaning if we accept a dualist theory of perception: physicalist qualities are objective, while phenomenal qualities, qualities associated with the experiences, concerns or emotions of human beings, are subjective. But if we reject dualism, the above distinction becomes wholly unclear.

---

[1] J. J. C. Smart, op. cit. p. 149

I suggest that by objective we mean *inter-personal*, and by subjective we mean *personal*, that which is not inter-personal. Thus colours, relative to a group of non-colour-blind people, are objective, but relative to all human beings, or, even more generally, relative to all rational beings, are subjective. It should be noted that in calling a quality either objective or subjective we are making an implicit appeal to a group of people. It should also be noted that in calling a *quality* subjective we do not imply that the quality does not really exist (or does not exist 'independently of the observer'), and in calling a *description* subjective we do not imply that the description is not really true. We imply only that not all members of the group of people in question can perceive the quality, fully understand the description.

Given this definition of objectivity, we may grant the physicalist that his is the only description that is objective relative to all rational beings.

University of Manchester

CONFIRMATORY MODELS OF THEORIES *

ROBERT ACKERMANN

THEORIES have often been construed by philosophers of science as deductive axiomatic systems at least some of whose theorems may be interpreted as expressing claims whose truth or falsity may be determined by appropriate scientific observation. At one or more steps of the derivation of such theorems, a correspondence rule may be used to derive some statement containing only observational terms from one or more statements containing at least one essential occurrence of a theoretical term. Theories of confirmation associated with construals of theories as deductive axiomatic systems have usually assumed that it is possible to determine whether or not a relationship of derivability holds between the axioms of a theory and any of various statements whose truth or falsity can be determined by observation. A number of cogent objections can be brought against the deductive axiomatic construal of theories (hereafter called the axiomatic system construal) and the theories of confirmation which are compatible with this construal. In this paper I should like to review briefly a few of these objections and then describe an alternative construal of theories and a theory of confirmation to accompany it that seem to avoid these objections in so far as the description of the construal can be made precise in this context.

There are some objections to the axiomatic system construal that may possibly be met by placing more stringent syntactical and semantical restrictions on those axiomatic systems which are considered as possible theories. Some philosophers have pointed out that the criteria of meaning for theoretical terms so far proposed in connection with this construal of theories do not result in distinguishing presumably useful theoretical terms from nonsense terms or unwanted metaphysical terms which might occur syntactically in the axioms of an axiomatic system.[1] The axiomatic system construal, since it fixes a theory as an axiom set and its consequences, may not even be compatible with any

---

* Received 24.vi.65

[1] For an examination of arguments related to this point see Peter Achinstein, 'Theoretical Terms and Partial Interpretation', this *Journal*, 1963, **14**, 89–105. I should like to thank the referees of my paper for several important suggestions toward improving its line of argument.

explanation of the generally accepted importance of theoretical terms in scientific theory. Indeed, if an axiomatic system T construed as a theory containing both theoretical terms and observational terms has certain syntactical features, then the set of all the theorems of T whose truth values may be determined by observation (and which consequently contain only observational terms as extralogical vocabulary) is identical with the set of theorems of another deductive system T^1 whose axioms contain none of the theoretical terms of T. Apparently, if a theory is defined as an axiomatic system, its observational content can be captured by another axiomatic system which does not contain any theoretical terms.[1] As a result of these two objections, it does not appear at the present time that the axiomatic system construal of theories can provide any sound account of the (acknowledged) historical importance of theoretical terms, nor can it provide a satisfactory means of differentiating between crank proposals and sound theories both of which could be expressed as axiomatic systems.

In addition to these technical objections, the construal of theories as axiomatic systems has been attacked by a wide range of objections designed to show that this construal is not compatible with same salient characteristics of scientific practice. Most of these objections centre about some claim that when scientists talk about or work with some theory, their talk cannot be construed as talk involving an implicit axiomatic system, since such a system is determined by its axioms and rules of inference, and yet a scientist will *revise a theory* and continue to regard it as the same theory even though rational reconstruction of his talk would involve first one, and then another, axiomatic system. In short, the axiomatic system construal can provide no reasonable account of how scientists may consider themselves to be revising and expanding a single theory over a period of time without doing considerable violence in reconstruction to what scientists say during their activities. The reluctance to give up the axiomatic system construal on the part of many philosophers of science is no doubt related to the fact that these objections have not led to an alternative construal of theories that enables a philosopher to retain some formal control over his material. I should now like to describe an alternative account of theories which seems to provide formal control while being compatible with the seemingly organic growth of theories.

---

[1] An extended discussion of this difficulty appears in Carl G. Hempel, 'The Theoretician's Dilemma', in H. Feigl, M. Scriven and G. Maxwell (eds.), *Minnesota Studies in the Philosophy of Science*, vol. II, Minneapolis, 1956

The view I would like to describe will be called the *two languages construal of theories* and it involves a deductive system, a syntactical explanation relation, an observational language with an associated theory of measurement, and a transformation relationship between certain statements of the observational language and a (proper) subset of the theorems of the deductive system.[1] In what follows, I shall assume that a satisfactorily clarified notion of explanation is available to us, and that this notion embodies the (somewhat controversial) claim that any statement B explained by another statement A is a deductive consequence of A. Hempel's work has shown that the deductivity requirement, even if necessary, is not sufficient, but that if it is taken as a necessary condition, certain syntactical restrictions may be developed for given languages as formal necessary criteria for a deductive argument to be an explanation.[2] Whenever a statement B may be deduced from a statement A in a language and all of the formal necessary criteria for explanation in that language are met, we will say that *A potentially explains B.*

The observational language of a theory will be a language in which various scientific observations of some kind of phenomena can be expressed, and in which various generalisations about sentences reporting these observations can be expressed and confirmed. An observational language may be a very restricted fragment of English, just wide enough to report various observations and allow generalisations on them to be formulated without being so wide that it contains observational reports about phenomena other than those being principally studied. In other words, the observational language is such that it is reasonable to suppose that if it were formalised, its extralogical primitive vocabulary would consist solely of *observational predicates* in the intuitive sense relied upon by adherents of the partial interpretation construal of theories. The logical vocabulary of the observational language will not include or permit the identification of such predicates or functors as +, ×, and < in their usual mathematical sense since

---

[1] This view is similar to some other views involving two languages but differs from them in the use of a notion of explanation and in denying that the relevant observational generalisations can be *deduced* from theoretical axioms. Compare Rudolf Carnap, 'The Methodological Character of Theoretical Concepts', in H. Feigl and M. Scriven (eds.), *Minnesota Studies in the Philosophy of Science*, vol. 1, Minneapolis, 1956, and Wilfred Sellars, 'Theoretical Explanation', in *Philosophy of Science: The Delaware Seminar*, vol. 2, Bernard Baumrin (ed.), New York, 1963

[2] Syntactical restrictions are developed in Carl G. Hempel and Paul Oppenheim, 'Studies in the Logic of Explanation', *Philosophy of Science*, 1948, **15**, 135-175

$(x)(Sx \supset Ux)$ must be a theorem of the deductive system if they are to stand in the confirmatory model relationship. As an observational language and a deductive system may not have the same underlying logic or underlying grammar, a full definition would require us to stipulate that one condition that an observational language and a deductive system stand in the confirmatory model relationship is that *if* the statements of the observational language are expressed in the same underlying logic or grammar of the deductive system, then a correlation of extralogical vocabulary must be defined such that every statement of the relation structure of a simple observational language is transformed by the correlation into an isomorphic theorem of the deductive system.

Let R be a sentence expressing the relational structure of the observational language. (R might be the conjunction of all of the well confirmed generalisations of the observational language or any sentence equivalent to such a conjunction.) Suppose that the correlation above transforms R into a theorem R' of the deductive system, and let T represent the conjunction of the axioms of the deductive system. The second condition of the confirmatory model relationship is that T should potentially explain R' in the deductive system, that is, that T and R' satisfy the formal necessary criteria for an explanation in the language of the deductive system. It is this feature that prohibits the construction of trivial axiom sets like $\{T, T \supset L\}$ to define deductive systems standing in the confirmat [illegible] relationship to an arbitrary relation structure L.[1] Let [illegible] of the observational language be transform[illegible] the axiomatic system by the corre[illegible] ndition and the important properties of a scientific theory that have been brought into focus by Popperian philosophy of science.

At this point, so long as only simple observational languages are considered in a study of theories, there may seem a rather slender basis for choice between the two languages construal of theories and the construal of theories as deductive axiomatic systems. In fact, an epistemological point traditional in the literature since Descartes seems to tell against the axiomatic system construal even at this point. The point is this: in so far as a useful theoretical notion is an abstraction and conceptual simplification of some observable property or observable individual, the theoretical notion and the observable property or observable individual will have quite different semantical properties. To choose an example of Descartes's, *sun* as a theoretical term in astronomy may designate a sphere, whereas *sun* as an observation term may designate only a spherical object, or perhaps a bright elliptical disk. If a deductive axiomatic system with an astronomical interpretation contains the name *sun* as a theoretical term in its axioms, then it is illegitimate to construe any theorem following from these axioms using the same *sun* as an observation claim about the sun construed as an observable object, since *sun* would then have shifted in its semantical properties, contrary to the restrictions placed on sound deductive inference. This argument suggests that the relationship between theory and observation cannot be that of the axioms of an axiomatic system to various theorems [illegible]stem whose truth is open to direct inspection.

An approp[illegible]dence rule can mitigate this difficult[illegible]strual of theories is

$(x)(Sx \supset Ux)$ must be a theorem of the deductive system if they are to stand in the confirmatory model relationship. As an observational language and a deductive system may not have the same underlying logic or underlying grammar, a full definition would require us to stipulate that one condition that an observational language and a deductive system stand in the confirmatory model relationship is that *if* the statements of the observational language are expressed in the same underlying logic or grammar of the deductive system, then a correlation of extralogical vocabulary must be defined such that every statement of the relation structure of a simple observational language is transformed by the correlation into an isomorphic theorem of the deductive system.

Let R be a sentence expressing the relational structure of the observational language. (R might be the conjunction of all of the well confirmed generalisations of the observational language or any sentence equivalent to such a conjunction.) Suppose that the correlation above transforms R into a theorem R' of the deductive system, and let T represent the conjunction of the axioms of the deductive system. The second condition of the confirmatory model relationship is that T should potentially explain R' in the deductive system, that is, that T and R' satisfy the formal necessary criteria for an explanation in the language of the deductive system. It is this feature that prohibits the construction of trivial axiom sets like $\{T, T \supset L\}$ to define deductive systems standing in the confirmatory model relationship to an arbitrary relation structure L.[1] Let statements A and B of the observational language be transformed into statements C and D of the axiomatic system by the correlation of vocabulary given in the first condition and suppose that C potentially explains D in the deductive system. The third condition of the confirmatory model relationship is that if C potentially explains D under these circumstances, then A must explain B in the observational language. A theory is then construed as the *claim* that a given confirmatory model relationship between an observational language and a deductive system will continue to hold as new additions to the relational structure of the observational language are made by means of new observations and generalisations on them. Since a theory is now taken as a claim that a given confirmatory model will continue to hold, then a theory on the two languages construal is a *falsifiable* claim that cannot be proved true. This means that it has the important properties of a scientific theory that have been brought into focus by Popperian philosophy of science.

---

[1] $\{T, T \supset L\}$ is not an explanans set for L on any of the extant formal models of explanation

At this point, so long as only simple observational languages are considered in a study of theories, there may seem a rather slender basis for choice between the two languages construal of theories and the construal of theories as deductive axiomatic systems. In fact, an epistemological point traditional in the literature since Descartes seems to tell against the axiomatic system construal even at this point. The point is this: in so far as a useful theoretical notion is an abstraction and conceptual simplification of some observable property or observable individual, the theoretical notion and the observable property or observable individual will have quite different semantical properties. To choose an example of Descartes's, *sun* as a theoretical term in astronomy may designate a sphere, whereas *sun* as an observation term may designate only a spherical object, or perhaps a bright elliptical disk. If a deductive axiomatic system with an astronomical interpretation contains the name *sun* as a theoretical term in its axioms, then it is illegitimate to construe any theorem following from these axioms using the same *sun* as an observation claim about the sun construed as an observable object, since *sun* would then have shifted in its semantical properties, contrary to the restrictions placed on sound deductive inference. This argument suggests that the relationship between theory and observation cannot be that of the axioms of an axiomatic system to various theorems of the system whose truth is open to direct inspection.

An appropriate use of one kind of correspondence rule can mitigate this difficulty, but even if the axiomatic system construal of theories is augmented by appeal to correspondence rules, this construal of theories must ultimately rely on some notion of *experimental error* to explain away discrepancies between theoretical predictions and measurements. But *experimental error*, unless receiving some reasonable explication, is an insufficient notion for distinguishing between random observations of a value which follows from theoretical axioms and random observations of a value not deducible from those axioms but closer to it than the range of possible error in the observations. The difficulty is not simply that of relating a precise theoretical claim to an imprecise observational claim, since it is the case that under some circumstances a precise theoretical claim may be falsified by a substantiated but somewhat imprecise observational claim. The difficulty is exactly that of explaining the circumstances which [illegible] one to say that a precise theoretical claim is confirmed by an observational measurement *no matter how precise.*

Utilising an astronomical example once more, let us suppose that a consequence of astronomical theory is that the sun and the earth are exactly 93,000,000 miles apart at time *t* and suppose that a best series of triangulations shows that the sun and the earth are closer than 93,000,002 miles and farther than 92,999,998 miles apart at time *t*. Measurements will in general have the property of giving the value of some metrical predicate within a certain range. A working scientist may rightly consider the triangulation measurement as confirming the theory. As a first approximation (which is all that will be considered in this paper) he may explain his reason for this as the fact that the theoretical distance is *within the range* of distances compatible with the observations. If '*Dt* (*sun*, *earth*) = 93,000,000 *miles*' expresses a statement of the appropriate theoretical language (or deductive system) and *Dt′* (*sun′*, *earth′*) = (92,999,998; 93,000,002) *miles′*' a statement of the appropriate observational language, then under the usual restriction of similarity of grammar and logic, the language replacement correlation for simple observational languages could be carried out with the additional stipulation that the range of measurement of *Dt′* (*sun′*, *earth′*) must be some interval of *miles′* which includes the theoretical number of *miles* given by *Dt* (*sun*, *earth*).

In general, suppose that an observational language is like a simple observational language except that it contains in addition at least some measurement ranges. Suppose also that the logic and grammar of the observational language and some deductive system are such that the relational structure of the observational language may be transformed into theorems of the deductive system by a vocabulary correlation as above with the additional stipulation that if some statement of measurement in the observational core of the observational language translates into a theorem of the deductive system under the correlation except for the range of measurement and some value given in the theorem, then the range must transform into the value. The transformation of a range to an (ideal) precise number in the deductive system is allowed if and only if the precise number is one of the values of the measurement range. An immediate difficulty is that a (standard) statistical theory associated with an observational language will assign a probability value *r* to the measurement range to indicate the probability that the true value lies within it. Any transformation such as the above could be carried out in its measurement range aspect by allowing *r* to approach

0. To prevent such trivial transformations, it is stipulated that the r associated with any transformation of a measurement range be 0·5 or greater. With this stipulation, the two languages construal may be generalised to include the case where an observational language includes measurement ranges. It is important to notice that this is not equivalent to adopting a correspondence rule for such cases, as a given measurement range may occur more than once, but need not be correlated to the same (theoretical) value in all of its occurreences, but only to a value within its range. As a further generalisation, deductive systems including statistical theorems could have confirmatory models whose derived statistics were ranges which would include the correlated statistics of the deductive system. Each of these two latter kinds of confirmatory models will have a probability value r ($r > \cdot 5$) associated with each transformation involving any statistical measurement theory. The lowest value of any such r associated with some transformation of a confirmatory model will be called the *least probability value* of that model.

One of the most interesting *prima facie* characteristics arguing for the two languages construal is related to the somewhat complicated notion of confirmation that must be developed to accompany it. Given the three conditions of the confirmatory model relationship, a theory can be said to be confirmed under a variety of circumstances.

If a generalisation which is not a logical consequence of the existing relation structure is added to the relation structure of an observational language through observation, then it must be transformed into a theorem of the deductive system under the conditions of the confirmatory model relationship or the theory is disconfirmed. Further confirmation may be obtained when a new generalisation is involved in explanations in the observational language which transform into potential explanations in the deductive system. Explanation relationships among the statements of the observational language should be transformed into potential explanations of the deductive system. Failure of such transformations to hold disconfirms the theory. Let us call the first kind of confirmation where a new confirmed generalisation of the observational language is added to its relation structure and found to be transformed into a potentially explained theorem of the deductive system *sentence confirmation.* Confirmation resulting from the introduction of a new explanation to the observational language mirrored in a potential explanation of the deductive system will be called *explanatory confirmation.* Any explanatory confirmation is also a sentence deducible from the theorems of the deductive system, as is often the case on the axiomatic system construal of theories.

The two languages construal has potentially interesting consequences for other kinds of problematic confirmation. To take one example, philosophers of science have often been puzzled by the apparently contradictory behaviour of a scientist who takes measurements to confirm one theory using an instrument whose workings he explains in terms of generalisations seemingly incompatible with the theory being tested. A person who believes that light rays do not have rectilinear paths but makes astronomical observations with a telescope whose workings he explains in terms of rectilinear paths seems to such philosophers to be contradicting himself. Indeed, it has been suggested in this connection that a telescope cannot be used to make observations unless a theory of light paths is already accepted to explain its workings. On any partial interpretation of an axiomatic system construal, the derivation of observational generalisations from the theoretical axioms means that any other course will involve the observer in the simultaneous acceptance of logically contradictory claims about the paths of light rays. Now since the account of an observational instrument's workings can at least in principle be reduced to observational generalisations and their consequences (nothing theoretical is put together to make an instrument), such generalisations may be part of the theory of measurement of the observational lang[illegible]ithout entailing a contradiction in conjunction with some th[illegible]he deductive system. If we have convinced oursel[illegible]tion of the telescope's workings that what we [illegible]relation to various scientific generalisations would then be seen as the result of a generalisation belonging ambiguously to each of two languages if its vocabulary happened to consist solely of extralogical terms which *as inscriptions* could be taken as belonging to either language.

A question of very great importance is the question of how one can choose between alternative deductive systems each of which may stand in the confirmatory model relationship to some given observational language. Using the notion that such deductive systems are blueprints or development plans, this may be phrased as the question of why we should expect an observational language to develop one way rather than another. And when the question is phrased this way, the expected answer would be that one plan of development is inductively better supported than another.

Such support has often been conceived for other construals of theories as support from models of theories.[1] In the usual model notion involved, a theory is said to have an observational model when the theory, construed as a partially interpreted deductive system, is isomorphic in an appropriately defined technical sense to an axiomatised observational language. It is not clear from historical examples that this kind of model relationship obtains very often. I should like to suggest the notion of *analogous* theories to bring a suitable model notion for discriminating among rival theories into the two languages construal. We can easily define isom[illegible]ms between deductive systems. Suppose that two observatio[illegible]es stand in the confirmatory model relationship to deductiv[illegible] isomorphic. If so, we may call the tw[illegible]es in that they are both claims [illegible] various scientific generalisations would then be seen as the result of a generalisation belonging ambiguously to each of two languages if its vocabulary happened to consist solely of extralogical terms which *as inscriptions* could be taken as belonging to either language.

A question of very great importance is the question of how one can choose between alternative deductive systems each of which may stand in the confirmatory model relationship to some given observational language. Using the notion that such deductive systems are blueprints or development plans, this may be phrased as the question of why we should expect an observational language to develop one way rather than another. And when the question is phrased this way, the expected answer would be that one plan of development is inductively better supported than another.

Such support has often been conceived for other construals of theories as support from models of theories.[1] In the usual model notion involved, a theory is said to have an observational model when the theory, construed as a partially interpreted deductive system, is isomorphic in an appropriately defined technical sense to an axiomatised observational language. It is not clear from historical examples that this kind of model relationship obtains very often. I should like to suggest the notion of *analogous* theories to bring a suitable model notion for discriminating among rival theories into the two languages construal. We can easily define isomorphisms between deductive systems. Suppose that two observational languages stand in the confirmatory model relationship to deductive systems which are isomorphic. If so, we may call the two theories involved *analogous* theories in that they are both claims that confirmatory model relationships will continue to hold between certain observational languages and deductive systems of the same form. This construal *does not require* that every sentence of the one observational language has an isomorph in the other observational language, a result seemingly in accord with the features of historical theories that have sometimes been invoked to guide choices between rival contemporary theories. Now if two theories are proposed which relate a given observational language to non-isomorphic deductive systems, but only one of these theories has an analogous theory which has been tested and accepted by scientists, then it seems a reasonable inductive inference to hold that the observational language is most likely to develop according to the plan of the deductive system which deducible from the theorems of the deductive system, as is often the case on the axiomatic system construal of theories.

---

[1] The importance of models for theory construction on traditional grounds is carefully discussed in Mary Hesse, *Models and Analogies in Science*, Sheed and Ward, 1963

The two languages construal has potentially interesting consequences for other kinds of problematic confirmation. To take one example, philosophers of science have often been puzzled by the apparently contradictory behaviour of a scientist who takes measurements to confirm one theory using an instrument whose workings he explains in terms of generalisations seemingly incompatible with the theory being tested. A person who believes that light rays do not have rectilinear paths but makes astronomical observations with a telescope whose workings he explains in terms of rectilinear paths seems to such philosophers to be contradicting himself. Indeed, it has been suggested in this connection that a telescope cannot be used to make observations unless a theory of light paths is already accepted to explain its workings. On any partial interpretation of an axiomatic system construal, the derivation of observational generalisations from the theoretical axioms means that any other course will involve the observer in the simultaneous acceptance of logically contradictory claims about the paths of light rays. Now since the account of an observational instrument's workings can at least in principle be reduced to observational generalisations and their consequences (nothing theoretical is put together to make an instrument), such generalisations may be part of the theory of measurement of the observational language without entailing a contradiction in conjunction with some theorem of the deductive system. If we have convinced ourselves by close inspection of the telescope's workings that what we see through it bears such and such a relation to what it is pointed at, then the fact that it is pointed at a star instead of some object in a neighbour's garden is of no consequence to our account of its workings.

The deductive system involved in a theory may be looked at as a blueprint or development plan for the observational language. It is consequently to be expected that certain terms, as *sun* and *earth* in the astronomical example, will play a rôle in both the deductive system and in the observational language. Their rôles will be somewhat different: postulated in the deductive system and determined by observation in the observational language. This view suggests that a given generalisation involving only vocabulary introduced to both the deductive system and the observational language may be construed as analytic in the deductive system and synthetic in the observational language. Some of the paradoxes associated with the putative truth values of is of a kind that has already described the successful development of other observational languages. Where both proposed theories have analogous theories, we could let the number and diversity of the theories analogous to the alternatives provide the basis for an inductive choice which may be more difficult than, but no different in principle from, any complex inductive decision. This slight shift would permit us to say vis-à-vis a much discussed example, ' What we know about atoms can be incorporated into a kind of theory that has proved successful in defining theories about the behaviour of billiard balls, or any other roughly spherical bodies which move in roughly straight lines and collide elastically with one another, and this looks like a promising line to develop '. On the more traditional model view, we are more or less supposed to feel that billiard balls are *like* atoms, because a fully interpreted billiard ball theory could be isomorphic to a theory about atoms, and that supposition has always seemed to many philosophers of science to be somewhat tenuous.

Where an inductive choice is not possible because both theories receive the same weight, or because neither has an analogous theory, other evidence may enable a choice between theories to be made. Simplicity is no doubt a weight where it may be invoked to discriminate between theories. An interesting method of justifying a choice in cases where analogous theories and simplicity fail to discriminate between rival theories is provided in some cases by the theory of measurement of the observational language. Let each of two inductively equivalent theories of similar simplicity be such that they stand in the confirmatory model relationship to an observational language. Now the theory of measurement of the observational language may have undergone a series of refinements in the past, and this series may have an interesting consequence. Suppose that the refinements in measurement are such that they tend toward a (posited) limit which is compatible with only one theory. One kind of refinement in measurement might narrow the confidence interval of a statistical relationship at a given level over a period of time such that, even though both deductive systems contain theorems compatible with the narrowest confirmed confidence interval (the interval with highest acceptable r) at the moment, an *expected* future confidence interval given the trend noticed in the intervals will be transformable into a theorem of only one of the two deductive systems. Under these circumstances, it seems reasonable to take the theory utilising this deductive system as the more reasonable plan of the observational language's future development.

It is a commonplace historical observation that some refinements in measurement can destroy confirmatory model relationships. An often discussed example from the history of science is that certain well confirmed gas laws turned out to be incompatible with predictions from the laws when the behaviours of gases could be measured under new techniques greatly extending the previously observable ranges of temperature, volume, and pressure. In such a case, a trend previously well confirmed by measurement may be suddenly upset. On the two languages construal, the same *true* observational reports and generalisations on them may simply be said to be transformed into one set of theorems rather than another. It does not follow that the theory constructed for the simpler observational language must now be considered false. In taking some measurements to build a bookcase in my room, I may be taking observations which can be expressed within some simple observational language, and I may only be interested in calculating some further observations which may be expressed in the same language. To do this, I do not have to consider my room as a solid in a relativistic, non-Euclidean universe, since there may be a well-confirmed theory for my simple observational language which will enable me to make the appropriate calculations, and this in spite of the fact that another theory exists employing a vastly extended observational language for certain purposes of scientific theorising. On the two languages construal, I may use any theory for convenience which is well confirmed with respect to an observational language in which the observations I am interested in can be expressed. On the axiomatic system construal of theories, I should strictly speaking have to regard the use of a rejected theory as the employment of false sentences and invalid rules of inference. The two languages construal thus seems to me to preserve many of the virtues of instrumentalism without entailing the awkward view which holds that theories are neither true nor false.

Washington University

DISCUSSIONS

Estimated Utility and Corroboration

Isaac Levi has raised some interesting questions concerning the more or less continuing dispute between Carnapians and Popperites.[1] What is novel about Levi's discussion is that it is supposed to concentrate *not* on the 'issue separating Carnap from Popper' but on a closely related issue.[2] Whether or not the issue considered by Levi is in fact the one 'separating Carnap from Popper', his discussion surely merits further consideration. In the first part of this note, Levi's criticism of Popper's rule of maximising corroboration will be challenged. In the remaining part, we shall challenge his defence of Carnap's rule of maximising estimated utility.

I

The important philosophic question Levi raises is: When scientists are trying to determine whether or not a hypothesis is acceptable, 'ought [scientists] maximise probability, corroboration, or anything else?'[3] Levi argues that scientists ought *not* to maximise corroboration, but ought to maximise 'expected epistemic utilities'.[4] That is, Levi rejects Popper's principle

> (MC) Given a set of mutually incompatible hypotheses choose that hypothesis with the maximum degree of corroboration relative to the available evidence.[5]

but accepts Carnap's principle

> (MEU) Among the possible actions choose that one for which the estimate of the resulting utility is a maximum.[6]

Levi has two arguments against MC. His first argument is supposed to 'turn the tables on Popper's celebrated refutation of Carnap'.[7] There is no need to consider this argument in detail since Levi grants that it 'depends upon two important assumptions' *and* he agrees 'for the sake of the argument' that 'Popper is right' in rejecting the first of the two assumptions.[8]

---

[1] I. Levi, 'Corroboration and Rules of Acceptance', this *Journal* 1963, 13, 307-313.

[2] Ibid. p. 308.

[3] Ibid. Levi raises both a normative and a descriptive question but considers only the former.

[4] Part of the time Levi uses the term 'epistemic utilities' and part of the time simply 'utilities'. I have taken the same liberty.

[5] Levi, loc. cit. p. 307.

[6] Rudolf Carnap, *The Logical Foundations of Probability*, Chicago, 1950, p. 269. This rule is defended by Carnap in a number of publications after 1950. See, for example. his 'Replies and Systematic Expositions', *The Philosophy of Rudolf Carnap*, ed. P. A, Schilpp, La Salle, Illinois, 1963, pp. 972 and 986; and 'The Aim of Inductive Logic', *Logic, Methodology and Philosophy of Science*, ed. E. Nagel, P. Suppes and A. Tarski, Stanford, California, 1962, p. 304.

[7] Levi, loc. cit. p. 308.

[8] Ibid. p. 309.

Levi's second argument then, is apparently the only one he intends to be decisive against MC. The argument runs as follows. First, assume

(*b*) a theory of rational 'epistemic' preference based on von Neumann-Morgenstern type postulates.[1]

These postulates entail MEU.[2] Thus MC is acceptable if, and only if, MC is equivalent to MEU. Call the sentence 'MC = MEU' '(A)'.[3] Now, a second consequence of the 'von Neumann-Morgenstern type postulates' (NM postulates) is

(I) No rational or scientific assignment of epistemic utilities must require that utilities vary with the available evidence.[4]

But Levi shows 'that if (A) is adopted, the utility functions generated do violate (I)'.[5] Thus (A) must be rejected and therefore, so must MC. This completes the outline of Levi's argument against MC.

My first objection to Levi's argument is that it is plainly question-begging. We begin with the question 'ought [scientists] to maximise probability, corroboration, or anything else?' A rule such as MC is evidently one of a number of alternative answers to this question. MEU is another alternative. Levi's first assumption (*b*) entails MEU. Thus the first premise of Levi's argument against MC entails the major alternative to MC. Hence Levi has no choice but to conclude that MC is an acceptable alternative if, and only if, it is equivalent to the alternative already forced upon us, namely, MEU. That is, granting Levi's assumption (*b*) is tantamount to granting equivalence to MEU *as a condition of acceptability* for any answer to the question 'ought [scientists] to maximise probability, corroboration, or anything else?' Surely this is an unacceptable *petitio principii*. A defender of MC would hardly allow Levi to determine which of the alternative rules is acceptable by assuming equivalence to MEU as a condition of acceptability.

Levi has, I think, suggested his reply to this kind of objection. After adopting assumption (*b*) he writes, 'should Popper balk at this procedure, he is obligated to supply an alternative theory'.[6] But curiously he rejects Popper's 'desiderata' as an alternative set of postulates to be satisfied. According to Levi, the fact that MC satisfies these 'desiderata' 'in no wise supports [Popper's] thesis that we ought to maximise corroboration any more than one can justify maximising expected utility by showing that expected utilities satisfy the postulates of an algebra of expectations.'[7] One would suppose, however, that the fact that MEU satisfies the NM postulates can not justify MEU *as far as any Popperite is concerned* any more than the fact that MC satisfies Popper's 'desiderata' can justify MC *as far as any Levite is concerned*. Why, after all, should the NM postulates be preferred as conditions of adequacy or as acceptability over Popper's 'desiderata?' Levi does not answer this question directly, but he seems to suggest that 'current theories of rational preference' support the use of these postulates (and consequently MEU).[8] This is reminiscent of Y. Bar-Hillel's attack on Popper's corroboration function from the point of view of the 'current theory of confirmation', i.e. from the point of view of the *prima facie* major alternative theory constructed by Carnap.[1] But the appeal to 'current theories' cannot be very convincing to an assailant of these theories.[2]

---

[1] Levi, loc. cit. p. 310.

[2] The proof of this is given in the Appendix of John von Neumann and Oskar Morgenstern, *Theory of Games and Economic Behaviour*, Princeton, 1947.

[3] In Levi's note '(A)' designates a proposition asserting the identity of Popper's corroboration function and Levi's estimated utility function. The rules concerning these functions have been substituted here to simplify the presentation. See, Levi, loc. cit. p. 310.

[4] Ibid. p. 311. [5] Ibid. [6] Ibid. p. 310. [7] Ibid. [8] Ibid. p. 313.

There is, I believe, a fairly plausible explanation of Levi's *rejection* of the agreement of MC with Popper's 'desiderata' as a justification of MC and his *acceptance* of the agreement of MEU with the NM postulates as a justification of MEU. Levi apparently believes that the NM postulates specify conditions of rational behaviour while Popper's 'desiderata' *only* specify conditions of an adequate corroboration function. But here we are being deceived by the phrase 'conditions of rational behaviour'. A person's behaviour is rational so far as it tends to satisfy the person's needs, goals or purposes. Hence, if the NM postulates may be used to specify rational behaviour, it is because action in accordance with these postulates can satisfy certain selected needs or purposes. Similarly, Popper's 'desiderata' may be used to specify rational behaviour *for scientists*. That is, a Popperite would argue that given the particular needs, purposes and interests of scientists (*qua* scientists), it would be irrational to appraise hypotheses on the basis of some function that does not satisfy Popper's 'desiderata'. In this sense, Popper's 'desiderata' may be used to specify rational behaviour. Therefore, although (as is shown in part II) Popper's justification of MC does *not*, as a matter of fact, consist of merely showing that MC satisfies his more or less intuitively acceptable 'desiderata', it must be granted that such a procedure would be as legitimate for Popper as Levi's corresponding procedure.

It is not, of course, necessary 'to supply an alternative theory' in order to give sufficient reasons for rejecting Levi's theory. The progress of science and philosophy would be considerably curtailed by such a requirement. It is one thing to show that a theory contains empirically false or contradictory propositions and quite another to invent a new theory. Most of us can do the former, but can *not* do the latter. Let us then, consider two problems of the NM postulates.

In the first place, it is well-known that one of the NM postulates, namely, the postulate of the transitivity of preferences, is false. That is, von Neumann and Morgenstern assume that if a rational being prefers x to y and y to z, then he prefers x to z. But such 'circular triples', to use Kenneth May's phrase,[3] as the following show that this assumption is false and raise serious doubts concerning the transitivity of preferences as a postulate of rationality.

> A preference cycle said to have been observed during the war concerns the behaviour of pilots in burning planes. When confronted with desperate choices between pairs from the set{flames, red hot metal, falling,} pilots most often made the choices (flames) P [is preferred to] (red hot metal), (red hot metal) P (falling), and (falling) P (flames).[4]

To avoid falsification of the transitivity postulate by such cases as the above, Levi *could* simply deny that the pilots (for example) were behaving rationally. Given such

---

[1] Y. Bar-Hillel, 'Comments on "Degree of Confirmation" by Professor K. R. Popper', this *Journal*, 1955, **6**, 157.

[2] Popper's reply, for example, was simply, '*there is no "current theory of confirmation"*'. Karl Popper, '"Content" and "Degree of Confirmation": A Reply to Dr. Bar-Hillel', this *Journal*, 1955, **6**, 158.

[3] Kenneth O. May, 'Intransitivity, Utility, and the Aggregation of Preference Patterns', *Econometrica*, 1954, **22**, 5.

[4] Ibid. p. 7.

Y

extraordinary alternatives, one does not expect and indeed one does no behaviour. However, the most cursory inspection of some of the designed to test the transitivity postulate rapidly convinces one that ' cir are not merely a product of peculiar circumstances.[1] In view of the amo mental data gathered against the postulate, those who continue to affirm consider it part of the *definition* of ' rational behaviour '. But, as May question then is whether rational behaviour as so defined has very much either descriptive or normative '.[2]

In the second place, the adoption of the NM postulates makes the res theory ' useless as a *prescriptive* [i.e. normative] theory of risky choi added). A second assumption of von Neumann and Morgenstern is that always determine which of *any* two risky outcomes he prefers. That mea can be used to select the hypothesis one ought to select *only if* one is ab determine which of any pair one ought to select. ' It is clear, then,' wri ' that in the context of decision under risk, the von Neumann-Morge theory is useless. It assumes as given what, according to the theory of d risk, is supposed to be calculated from *other* givens.'[4]

II

The results of part I show that (1) the first of Levi's two arguments a abandoned; (2) the second begs the question; and (3) the NM p questionable.

As Levi presents his case against MC it is apparent that even if the N are rejected, he shall assume postulates which ' at the very least . . . justif expected epistemic utilities '.[5] That is. MEU seems to have some kind status *a priori*. The remainder of this part of our discussion consists of Levi's defence of MEU.

Levi summarises his defence of MEU as follows.

> [MEU] takes into account Popper's expressly stated views concerning importance of hypotheses while weighting such importance with pr function the maximisation of which is justifiable according to curre rational preference.[6]

So far as the second half of this claim is concerned (i.e. from ' while ' on) o have already been stated. The assertion that MEU ' takes into acco expressly stated views concerning the scientific importance of hypotheses' is gratuitous. In fact, MEU fails exactly where Popper claims Carnap's explication of ' degree of confirmation ' fails, namely, where unrestricted universal hypotheses are involved.[1] In Levi's own words,

---

[1] The most famous ' circular triple ' is the ' paradox of voting '. This a described by May in the third section of his article.

[2] Ibid. p. 8. Warren McCulloch summarises the objections to the transi ate as follows. ' Experimental aesthetics, economics, and conditioned refl produced instances in which, under constant conditions, preference was c such instance would have been sufficient basis for categorical denial of th that values were magnitudes of any one kind. Thus for values there can be scale.' ' A Heterarchy of Values Determined by the Topology of Ne *Bull. Math. Biophysics*, 1945, 7, 92–93.

[3] Anatol Rapoport, *Strategy and Conscience*, New York, 1964, p. 20. If be made which leads to one of a set of possible outcomes whose individual are known, the choice is said to be ' risky '.

[4] Ibid. p. 21. R. Duncan Luce and Howard Raiffa in *Games and D* York, 1964, pp. 31–32 evidently support Rapoport's view.

[5] Levi, loc. cit. p. 309. [6] Ibid. p. 313.

> In general, the use of the Bayes method with epistemic utilities when choosing among universal hypotheses will result in identifying the expected utility of accepting a universal hypothesis with the utility of accepting such a hypothesis when it is false no matter how we assign these utilities.[2]

He grants, therefore, that MC ' enables us to choose between hypotheses in important cases where [MEU] fails utterly '.[3] But Levi offers three replies to those who would object to MEU on this ground.

(1) It is possible to find a ' probability measure that does permit probabilities greater than O for universal hypotheses '.[4] Carnap's theorems of ' instance ' and ' qualified-instance confirmation ', for example, are designed to do just that. Unfortunately, Carnap's two theorems have been severely criticised.[5] Indeed, while it must be granted that such a ' probability measure ' *might* be constructed, we have yet to see a satisfactory one.

(2) MC still requires ' justification regardless of the merits of the Bayes method as a tool for criticism '.[6] Now, Popper's justification of MC has never taken the form of Levi's justification of MEU. That is, Popper has never claimed that his ' desiderata ' are merely more or less intuitively acceptable and satisfied by MC. He has, however, repeatedly offered one argument since 1934 in defence of MC which seems to have been systematically avoided by most philosophers of science.[7] This argument runs as follows. One of the primary goals of a scientist is the discovery of laws or theories that adequately describe the world in which we live. The elimination of a number of logically possible but erroneous descriptions of ' our particular world ' is a necessary condition of determining an adequate description of ' our particular world '. Of all the hypotheses we have, those with the highest degree of corroboration (and therefore the lowest initial or relative probability) contribute most toward the elimination of inadequate but logically possible descriptions of our world. Therefore of all the hypotheses we have, those with the highest degree of corroboration are the most acceptable. This is Popper's primary justification of MC.

(3) Following Popper, a defender of MEU might appeal to some ' fine structure measures ' to obtain more acceptable utility values.[8] Instead of attempting a thorough discussion of this suggestion, we wish to point out that although it appears to be only an *ad hoc* solution for *both* Popper and Levi, one's judgment should be suspended until the details of the solution are presented.

State University College
Plattsburgh, New York

ALEX C. MICHALOS

---

[1] See, for example, Popper, loc. cit. p. 158.

[2] Levi, loc. cit. p. 312.

[3] Ibid.

[4] Ibid. p. 313.

[5] See, for example, Popper, loc. cit. pp. 160–161 and Alex C. Michalos, ' Two Theorems of Degree of Confirmation ', *Ratio* (to be published).

[6] Levi, loc. cit.

[7] The argument appears first in Karl Popper, *The Logic of Scientific Discovery*, London, 1959, pp. 113 and 270. It reappears in Popper's ' Degree of Confirmation ', this *Journal*, 1954, 6, 146, and *Conjectures and Refutations*, New York, 1962, pp. 286–287.

[8] Levi, loc. cit.

KARL R. POPPER

A NOTE ON THE DIFFERENCE BETWEEN THE LORENTZ-FITZGERALD CONTRACTION AND THE EINSTEIN CONTRACTION

LORENTZ and Fitzgerald thought that the contraction of measuring rods is caused by the movement of matter through the ether, and that the contraction, though not measurable within the moving system (because it affects *all* matter and *all* measuring rods equally) is a physical effect exerted by the ether upon moving matter. (This would be understandable if ' we assume that molecular forces are also transmitted through the ether, like the electrical and magnetic forces ', as Lorentz put it in 1895; in 1904 he assumed that the electrons are deformed.) Moreover, they do *not* say that a rod *at rest in the ether* would *appear contracted* if measured with the rods of the moving system. (Rather, one might expect, by implication, that it would appear *expanded.*)

Einstein suggested that the contraction is a question of ' perspective ' rather than a physical contraction in the moving matter, and that it is mutual. Also, if for some reason we should consider the system A at rest (say, for argument's sake, in the ether) and B moving with velocity v, the rods of the system B would look contracted not only if measured with the rods of A, but also, and *by exactly the same amount,* if measured with the rods of a system C which, relative to B, moves with the same velocity v with which B moves relative to A (so that, according to Einstein's rule for adding velocities, C will move relative to A with a velocity $u = 2v/(1+v^2/c^2)$ in the same direction as B). That is to say, we get the *same* contraction if we measure the B-rod in two systems, A and C, which move with a high velocity relative to each other.

It is this situation which shows that none of the (inertial) systems is to be preferred as being at rest, and which thus leads to the abandonment of the ether, that is to say, of the assumption that there is a system, such as A, which is at rest relative to the ether (and therefore ' absolutely ' at rest): it is impossible that A and C are both at rest, yet the rods in B behave symmetrically with respect to A and C. Or to put it in another way: the conjecture that the ether makes the rods of the moving system B contract relative to the system A (which was assumed to be at rest in the ether) is to be rejected according to Einstein, because the rods of B are contracted in exactly the same manner relative to C which cannot possibly be at rest in the ether if A is at rest in the ether.

These remarks have some bearing on the recent article in this journal, on ' The Origin and Concept of Relativity ' by G. H. Keswani.[1] The article contains much that is valuable; yet the author's scrupulous attempts at fairness and his obvious scholarly care do not prevent him from arriving at what I think is a mistaken assessment of Einstein's work. Though Einstein appears to have known Poincaré's *Science and Hypothesis* prior to 1905, there is no theory like Einstein's in this great book. Einstein could not have known Poincaré's article of 1905; and even in this article there is only a most inspiring programme sketched for a relativity theory—not the theory itself. Thus Poincaré did not, as Einstein did, show how the mysterious

---

[1] G. H. Keswani, ' On the Origin and Concept of Relativity ', this *Journal*, **15**, 286–306, and **16**, 19–32.

Lorentz-Fitzgerald contraction could be replaced by something completely different. In other words, though Poincaré has to be credited with predicting the coming of what is rightly called Einstein's theory of relativity, it was Einstein who achieved it.[1]

University of London KARL R. POPPER

---

[1] The expressions 'longitudinal mass' and 'transverse mass' are pre-Lorentzian, and the misprint in Einstein's formula for the transverse mass should not, in view of the correct equations that precede it, have been described as a mistake. Keswani's exclamation, in 16, p. 28, line 3, is unfounded and shows that he misunderstands the passage on which he is commenting. There are several similar points, e.g. p. 31, lines 18 to 20.

REVIEWS

SUBJECTIVE PROBABILITY

THE subjective theory on probability seems to solve the main problem about induction. That is, it both analyses the inferences which intelligent people draw from experience, and offers excellent reasons for thinking that those inferences are sound. It provides a foundation for inductive reasoning comparable to Aristotle's work in deductive logic and is already used by many investigators in drawing practical conclusions from complex experimental data. Among philosophers it has hitherto suffered from two accidents: the inaccessibility of its writings and the label 'subjective'. The present anthology[1] removes the former difficulty and also ends the latter, for once you see how 'subjective' is used in these essays you forget the word's stigmata.

The collection begins negatively. In 1866 a chapter of Venn's *Logic of Chance* attacks de Morgan, and for a long time no subjectivist can overcome Venn's dilemmas. Then in 1924 Borel reviews Keynes's *Treatise* and proposes personal betting rates as measures of probability. In 1926 Ramsey's 'Truth and Probability' finds a rigorous foundation for the idea. De Finetti's work, commenced about the same time, comes to a climax in 1937 with *La Prévision: ses lois logiques, ses sources subjectives;* he has added retrospective footnotes to the new translation in this volume. For sceptics who doubt his ready assumption that confidence in hypotheses can be measured, 'The Bases of Probability', written by Koopman in 1940, has a logic for comparing confidences without numerical measurement. Finally there is the post-war blossoming of the whole discipline, of which Savage gives some inkling as he looks back on it, and on his own contributions, from the vantage point of an address to the 1960 Berkeley Symposium on Probability and Statistics.

The editors have provided an excellent bibliography and a useful introduction. The book only seems expensive for a paperback of reprints, for unlike most others which have recently come to hand, it is newly laid out in clear type on good paper, and is altogether one of Wiley's finest productions. The mere act of collection has been a contribution to philosophy, for one can now make out two subjective theories. Hence this notice has two parts. (1) is on the conservative, or Bayesian, subjective theory, and (2) on a radical theory which threatens the whole traditional philosophy of science.

---

[1] *Studies in Subjective Probability.* Edited by Henry E. Kyburg, Jr., and Howard E. Smokler. John Wiley & Sons. New York and London, 1964. Pp. vii. + 203. Paperback, 36s.

1 *The Conservative Theory*

Intelligent people who at one time disagree about what is likely to happen, or about the truth of a scientific hypothesis, may later, after sharing some experiences, concur. They may even settle to their joint satisfaction what is true. They draw inferences from experience and call this learning from experience. They are not coerced into conforming with each other; at most their experiments 'force' them to agree. How *force*, if their observations do not entail what is agreed on? Hume said it is mere custom that we come to agree. The conservative theory has a different answer.

First it gives axioms for degrees of personal confidence. Any man can compare his degrees of confidence: 'I should be more confident in *h* if I knew *e*', he says, 'than of *i* if all I knew were *d*.' The modest axioms in Koopman's paper almost certainly fit such comparisons. A man may also quantify his confidence: 'Told *e*, I should be 60% sure of *h*.' Half a dozen strong arguments urge that if a man is sensible, his numerical degrees of confidence will satisfy axioms for the full probability calculus.

The most popular of these arguments, as first set out by de Finetti, equates a man's degree of confidence in *h* with the least odds at which he would bet on the truth of *h*. Now an avaricious man would be stupid if a Dutch book could be made against him—that is, if he were open to a system of bets such that he could not profit, while the bookmaker could. (Offer 5 to 1 on the defenders' keeping the America's cup at the next race, and 2 to 1 on their losing it; I can bet with you so you cannot win no matter what happens at sea, while I can win.) A set of betting rates avoids the possibility of a Dutch book if and only if it satisfies the probability axioms. It is inferred that the axioms provide a necessary, though far from sufficient, condition that a man's degrees of confidence are sensible.

More profound proofs, suggested by de Finetti, and completed by Koopman and by Savage in his book, *The Foundations of Statistics*, show how very feeble conditions guarantee that any metric imposed on comparisons of confidence will satisfy the axioms. Or you may follow Ramsey and get the axioms by jointly defining utility and degrees of confidence. Or avoiding utility you can adapt arguments from Jeffreys', *Theory of Probability* and from *A Theory of Probable Inference* by R. T. Cox. All this is to some extent old hat. In the paper posthumously sent to the Royal Society in 1763, Thomas Bayes established that degree of confidence—as measured by ratio of stake to prize in a betting situation—should satisfy the probability axioms. Indeed Bayes, like Jeffreys, may work with interpersonal and objective degrees of confidence, but his observations apply equally well to personal and entirely subjective degrees.

The next step is a trivial consequence of the axioms. Let *h* be a conjecture and *d* some initial data about *h*. Let *e* state the result of an experiment performed after *d* is known. We may speak of (i) your confidence in *h* given *d* alone: your *initial confidence* in *h*; (ii) what your confidence in *e* would be if you already knew both *h* and *d* but not *e*: this is the *likelihood* you attach to *e* in the light of *h*; (iii) your confidence in *h* after you learn of *e*: call this your *posterior confidence* in *h*. Bayes proved (iii) is proportional to (i) and (ii): if your degrees of confidence satisfy the probability axioms, your posterior confidence is in proportion to your initial confidence and the likelihood.

Changes in confidence represent the influence of experience. So Bayes's theorem tells how you ought to learn from the experimental data *e*. Better still, we prove that as you conduct experiments of the same kind, your degrees of confidence become increasingly stable. Best of all, take any two people with very different initial prejudices. So long as they agree on what is possible at all, and agree on the results of individual experiments, we prove that not only will their opinions become more stable, but their divergent views will increasingly coincide. When their degrees of confidence satisfy the axioms, experience dominates prejudice.

The Bayesian theory concludes that learning by experience (in the sense of the paragraph opening this discussion) is justified because (*a*) degrees of confidence ought to satisfy the probability axioms, and (*b*) people whose degrees of confidence satisfy the axioms, who make the same observations, and who agree on what is at all possible, will learn from experience in the required sense. Recent workers have checked that many people have degrees of confidence which pretty well satisfy the axioms. So learning by experience is not just a human custom, but a reasonable one. Having understood the custom, we can perfect it, using, for example, that fine introduction to practical applications: 'Bayesian statistical inference for psychological research', by Edwards, Lindemann and Savage. (*Psychological Review*, lxx (1963), 193-242.)

Those who disagree with all this will find most of their objections in Venn. His attack should be supplemented by Keynes's contention, in the *Treatise*, that very often different degrees of confidence are not really comparable. Koopman has more to say on that, and develops a logic weaker than de Finetti's which need not demand that degrees of confidence are always comparable; his system delivers a weak version of Bayes's results. So even if one is as sceptical as Keynes, much remains, and Borel's contribution to the present volume should do a good deal to dispel Keynes's doubt and to restore the full Bayesian theory.

Those who accept the full theory will find muddy waters beneath the smooth flow of theorems. Nearly every offhand phrase is full of difficulties. '*If people agree on what is at all possible*, their opinions will come to coincide.' The italicized qualification is essential, for Bayesian learning proceeds by multiplying initial degrees of confidence by likelihoods. If you think something impossible your confidence in it is 0, and no multiplication will alter that. Once you think something impossible, you will never change your mind without being inconsistent with the probability axioms. But often it is reasonable to be sure something is false and later come to think it might be true. Our theory cannot cover that sort of learning. Hence it must say either that you have non-zero confidence in any contingent proposition, or else that there is some kind of learning which is not Bayesian. The former alternative will appeal to sceptical philosophers, but of course it is wrong.

There is a question about the constant of proportionality mentioned in Bayes's theorem. It depends on the class of possible hypotheses, and if it is to be exact, each possible hypothesis must be precisely stated. (Cf. remarks by G. A. Barnard, p. 82 of *The Foundations of Statistical Inference* by Savage *et al.*) Hence the Bayesian theory must either arbitrarily limit the hypotheses which a person deems possible, or grant that evaluation of posterior confidence is not entirely exact. When practical statistics wants a definite model of the influence of data upon hypotheses, the former alternative is always used, but for the philosophy of induction only the latter is sound.

Then there are Goodman's riddles. 'Grue', it will be recalled, means examined before 1984 and green, or not examined before 1984, and blue. Contrast the hypothesis 'the frequency with which *Carnationus Wildeus* bloom green is between ·3 and ·4' and 'the frequency with which they bloom grue is between ·3 and ·4'. The former hypothesis may seem unlikely, but the latter is absurd. How do we know? From experience? If on our *tabula rasa* there were nothing to choose between the pair, no amount of Bayesian learning before 1984 could help. Goodman thinks this new riddle of induction is Hume's problem in disguise. If it is, the subjective theory offers no help for induction.

But the theory can contend that there are distinct problems: (1) the problem of justifying the ways in which we learn by experience which of the hypotheses which interest us are acceptable; (2) the problem of justifying our initial discrimination between hypotheses, and of counting some, like the one with 'grue', as altogether absurd. The theory can claim to solve the first problem, and it can be argued that there is no solution to the second: we can at most describe our habits, and characterise the hypotheses which we find interesting, say in terms of simplicity, maybe following Goodman's leads in that direction, or in statistics taking a subjective version of Jeffreys's simplicity postulate.

Much modern work implicitly distinguishes the two problems of induction. In Carnap's system the admissible hypotheses are those expressible in a simple formal language, and a confirmation function plots the path of learning. For Popper, you learn about hypotheses only by refuting them, and the admissible hypotheses are the refutable ones, graded by their 'content'. Neither theorist is safe against Goodman's riddles. Popper's measures of content collapse and a Goodmanic choice of primitive predicates embarrasses Carnap. The subjective theory has the unique merit, it seems, of making explicit how Hume's problem is not identical to Goodman's.

Humean difficulties fall in two. On the one hand the subjective theory answers Hume's own question about learning from experience. On the other, it grants that Goodman has a new and fundamental question, and replies to it in a Humean way.

2 *The Radical Theory*

If theories need forbears, count J. S. Mill among the ancestors of the most radical subjective theory. All reasoning is from particulars to particulars. Universal propositions are at most mnemonics or rules of inference; they do not state facts about the world. Or so Ramsey hesitantly urges in one of his last papers, ' General Propositions and Causality '. He saw the paradoxes in this view more clearly than did Mill, but had no time to resolve them. He was spurred on by his subjective theory about probability and by Brouwer's criticism of the use of universal propositions in classical mathematics. He would have been delighted by one tool forged shortly after his death, de Finetti's notion of *exchangeable events.*

De Finetti also thinks that ' the only events which interest us ' involve ' the occurrence of certain facts, more or less complex, but verifiable in a finite time '. He distrusts all species of universal proposition, but one galls him more than the others. Statistical hypotheses are commonly supposed to ascribe long run frequencies to some parts of the world, and hence to be about a dispositional property. De Finetti scoffs, saying they concern only ' a mysterious pseudo-property '.

Our conservative theory need not agree with de Finetti, for it can tolerate statistical hypotheses. The so-called subjective theory is a theory about personal degrees of confidence; there is nothing in its conservative version which precludes a person having confidence in a statistical hypothesis. But radical de Finetti thinks any such hypothesis is a fraud. Now when he mocks at long run frequency he has in mind antiquated analyses in terms of limits of an infinite sequence, and frequency theorists can do better now. So if his contribution were primarily a polemic against old work, it would be out of date, no matter how interesting its progeny. But we may regard it as truly radical; part of a scorched-earth policy in which all the traditional diet of philosophy of science—the maze of universal laws, general counterfactuals, dispositional predicates, statistical hypotheses—is consigned to the flames.

The traditional statistics subsists on the idea of independent trials of some kind on a chance set-up. What happens in practice? A florist wants to know the proportion of his next year's chrysanthemums which will breed true. He plants a dozen in his greenhouse. He accidentally steps on one soon after. The rest bloom in due course. Before even asking whether the trials are independent, we must establish what sort of trial has been performed. Twelve allegedly independent trials, with various possible results, including ' stepped on '? One trial consisting of twelve plantings, of which the eighth was injured by mildew or robins or sloppiness? Or a trial consisting of eleven 'real' plantings? De Finetti would ridicule this puzzle of Barnard's. The facts are that a florist has various degrees of confidence in the first plant blooming white, the second doing so, and so forth. Because of the way he chose his seeds he does not think the order of planting matters; when he began, he was as confident that his first plant should bloom white as that the 1000th from the same stock should do so when planted next spring. De Finetti gives a natural definition of this sort of personal indifference. When a person thinks some events satisfy this definition the events are called *exchangeable*, relative to his personal degrees of confidence. De Finetti proves that when a man thinks of events as exchangeable and conducts experiments he will form the same expectations about next year's crop as if he had postulated statistical independence, and long run frequencies, and used the Bayesian theory of inference. One had thought of statistical independence as a physical property of some part of the world, and as an essential core to statistical inference. De Finetti has proved it *de trop*. This is the first successful philosophical reduction in the whole turbid history of reductionism.

De Finetti thinks it is only the beginning. The study of exchangeability 'includes in reality the problem of reasoning by induction'. Moreover, 'everything that does not reduce to a simple statement, to an isolated historical truth, everything that councils us for the future ' is based on the laws of subjective probability and the analysis in terms of exchangeability. Hume asked, why am I so confident that if I eat this bread it will nourish me? De Finetti would answer that I conceive of many different eatings of bread, and have data about some of them; the outcomes of the different trials are exchangeable from my point of view; in virtue of exchangeability, of my positive knowledge, and of the probability axioms, I can derive great confidence that the bread before me is wholesome.

Universal laws and theories are missing from this account, and de Finetti has never examined them satisfactorily. Statistical hypotheses in practical inferences have been analysed away, but not universal statements; still less theories. How could de Finetti deal with their crucial role in the human mastery of the universe? We can at most return to the hints of Ramsey's last papers to guess how a radical opinion might go. Our present comprehension of the role of scientific theories is poor; radical work stemming from de Finetti and Ramsey may yet succeed in turning our preconceptions upside down. But this task has hardly been commenced. Savage, the most notable recent contributor to subjectivism, has primarily been a statistician. In statistics he inclines to de Finetti's radical views, although he is often content to speak as if he accepted the conservative theory, since that makes it easier to converse with colleagues. He has made the subjective statistics respectable among statisticians. Now we need a philosopher of comparable stature to make a radical theory on induction intelligible to philosophers.

IAN HACKING

Act and Agent: An Essay in Philosophical Anthropology. By Douglas Browning. University of Miami Press, Florida, 1964. Pp. 134. $3.95.

In this book Dr Browning, who is Associate Professor of Philosophy at the University of Miami, seeks (to quote the blurb on the dust-jacket) to lay 'the foundations for an original theory of man's nature' by providing 'a careful analysis of the nature of human action and agency'. Anthropology comes in because it is an attempt to answer the question 'what is man understood as a concretely encultured being?', and its first job is 'no other than the survey of all human concerns in an effort to determine their specific natures' (p. 122).

Few professional anthropologists today, in this country at any rate, would claim so all-embracing a field for their special study. It evidently takes in practically everything that can be thought about, so that although Professor Browning modestly tells us that the task of his essay 'may be conceived as a prolegomenon to any future anthropology which may be considered sound' (p. 3), his real concern is with 'the general problem of the nature of man' and, in particular, with the question how man can be conceived as a moral agent.

Thus the book is about the foundations of ethics rather than of anthropology, but this is not to say that it has no relevance for (social) anthropologists who, like all other social scientists, may sometimes ask whether the people who are their subject matter are 'free' moral agents (whatever this may mean) or mere automata, their behaviour causally determined and so in principle as predictable as that of a machine. However the more sophisticated among social scientists recognise that whatever the ethical import of this dilemma it is not important so far as scientific procedure is concerned: as Ernest Nagel (among others) has pointed out, behaviourism may be a method without necessarily being a metaphysic.

Browning, however, is not an empirical scientist but a philosopher, and for him the problem is about reality, not about method. His starting point is the inadequacy of the 'personality/behavior' approach—a kind of 'self-castration'. Philosophers and psychologists, he says, have mistakenly identified choice and behaviour, an error which has led to 'an unintelligible and unacceptable double vision of man as scientifically knowable yet unalterably insignificant, on the one hand, while perhaps morally significant but unalterably mysterious, on the other' (p. 2). Over and above behaviour, which is in time and space and is causally determined, there is 'a distinctive agent aspect of [man's] selfhood which is irreducible to personality', and the first of the three Parts into which Browning's book is divided is about this acting aspect of the self. Acts, or conscious choices, are *sui generis*; they are not reducible to behaviour nor susceptible of causal analysis. Choice is 'a brute fact of daily and mundane affairs'; but it is not a datum and it cannot be analysed as though it were.

Since there are acts there must be agents, and the concept of the agent is analysed in Part II. This concept, that of 'unified subjectivity' is necessary as 'the source and basis of the normative' and of enduring personal identity, but, again, the agent as subject, though indispensable, is not a datum but is 'by its nature forever hidden' (p. 94).

Part III, 'Reality and Man,' restates and develops some aspects of these themes, and suggests that the appropriate philosophical orientation in such matters is one of 'opacity'. 'We live in a universe which constantly tends to the creation of separate pools of opacity, secrecies into which no one may look, recesses hidden even to he [*sic*] in whose soul they lie' (p. 108). The author concludes that 'man is essentially non-datal, non-processional, and non-causal in nature', and he repeats (p. 118) that 'the universe is in some respects unalterably opaque'.

This summary does very much less than justice to Professor Browning's complex and often subtle argument. But to a reviewer who is not a professional philosopher it is not very clear in what way it represents any substantial advance on (rather than a restatement of) Kant's philosophy, to which Browning freely acknowledges his debt. It was, after all, Kant who showed that 'the transcendental unity of pure apperception' is a condition of experience, that the 'pure self' so implied (as opposed to the empirical self) could not be conceived as temporal or causally determined, and that moral agency (the good will) involves willing in accordance with norms (or 'maxims').

In the last resort (and Professor Browning, despite his initial scorn for behaviourist obscurantism, comes to see this in the end) the universe *is* opaque; we can never—at least by philosophising—get behind the intellectual constructs which we have contrived to deal with experience, and come to grips with 'reality' itself. If the 'agent' is by definition unknowable, the concept of it cannot help us very much in anthropological analysis, indispensable though it may be as the basis of morality. But for a reasonably clear and sometimes stimulating restatement of this perennial problem Professor Browning's book has much in it to interest philosophically minded social scientists.

On so important a theme, however, at least a select bibliography would be a help.

J. H. M. Beattie

Signification and Significance: a study of the relation of signs and values. By Charles Morris. M.I.T. Press, Cambridge, Mass., 1964. Pp. x+99. 47s.

For many years now, Dr Charles Morris has been writing on two related subjects: semiotic, the general theory of signs, and axiology, the general theory of values. This short book draws the two subjects together. Roughly speaking, it is divided into two parts, the first laying down the conceptual apparatus required for the second. Having presented his so-called 'model', he uses it to analyse certain theoretical and empirical work in the behavioural sciences. Admitting his debt to Mead's theory of action and behaviour, Morris touches on psychology, linguistics, and information theory. He also raises the question 'Is a work of art a sign?', and discusses at length the possibility of experimental studies in the relations between signs and values in the arts. The success or failure of all this clearly depends on the merits of the earlier half of the book. (And, of course, conversely: but it is occasionally possible to judge of the bluntness of a tool without using it.)

The 'basic terms of semiotic' are taken to be as follows:

> *sign*—the bee's dance is taken as an example;
> *interpreters*—in the example, other bees;
> *interpretants*—the bees' disposition to react in a certain way when confronted with the dance;
> *signification*—the kind of object towards which the bees are prepared to act in this way;
> *context*—this is notoriously difficult to explain, and in the chosen example Morris is content to remark that 'the position of the hive is part of the context'.

The underlying epistemology is simple enough—perhaps too simple. There are scores of points at which readers will wish to fly off at a philosophical tangent. For example, Morris does not wish 'to deny that we may signify objects present in immediate experience . . .', but has far too little to say on the 'objectivity' involved. Thus 'significations' in general are not to be regarded as entities 'in any objectionable sense'. They are rather 'certain describable aspects of complex behavioural processes in the natural world'. (This statement may well be reconcilable with what is said about the objects of immediate experience, but the book itself is of little help.) That the bees are disposed by the dance to seek food objects in a certain context is what we are said to observe. 'There is nothing "mythical" about significations when so interpreted', we are told. But once lulled into a sense of security, in the face of such a commonsense account, we are gently advised that significations are not all of one sort. Designative significations, namely those already mentioned, are 'stimulus properties of the object.' (Elsewhere, 'a sign is designative in so far as it signifies *observable* properties of the environment or the actor'.) An 'appraisive signification', however, although called a 'reinforcing property' of an object, is said to be a 'capacity of an object to increase the probability of the performance of a response made to it'. The word 'property' is therefore misleading, as Morris half realises. Thirdly, there are 'prescriptive significations' explained by the phrase 'act as instrumental'. It is not easy to see what is meant by this expression, although apparently such an act is one which results in a desired consequence. The main reason for speaking in this way seems to be to make comments about philosophical method or procedure. Bar-Hillel suggests that rational discussion between speculative and analytic philosophers is possible only if conducted in a metalanguage acceptable to the latter. Tennessen criticises the methods employed by some analytic philosophers for distinguishing between 'permissible' and 'impermissible' locutions and emphasizes the right of philosophers to make recommendations for changes in usage. Also concerned with methodology is Karl Menger's 'A Counterpart of Occam's Razor in Pure and Applied Mathematics; Ontological Uses', in which Menger proposes to supplement Occam's *Law of Parsimony* with a *Law against Miserliness*—'a maxim denouncing equivocations just as Occam's law opposes synonyms (p. 104). Menger discusses the term 'variable' as a case to which his law applies.

Evert W. Beth ('Extension and Intension') suggests an explication of the concepts of extension and intension with respect to subjects (assumed to be formalisable) such as Astronomy, Cosmology, or Chemistry. Let the subject be expressed in an elementary formal theory S. Then the extension of a constant term *t* of S with respect to a model M of S is the denotatum of *t* in M; and the intension of *t* with respect to a class X of models of S is the class whose members are denotata of *t* in some model in the class X. To get the intension of *t simpliciter*, Beth has the Class X of models correspond to all possible worlds. However, I cannot see how this latter move can successfully rule out non-standard models from the class X without circularity. For example, if Astronomy can be formalised in an elementary theory T, then T will have a model in which, say, 'Venus' denotes some number, and I cannot see how this pythagorean intrusion into the intension of 'Venus' could be eliminated short of saying that that model is not the sort *intended*.

W. V. Quine ('Carnap on Logical Truth') raises some of his well-known criticisms of the linguistic doctrine of logical truth. In the course of this he distinguishes between 'legislative' and 'discursive' postulation (p. 49). The latter 'involves mere selection from a preexisting body of truths' while the former 'institutes truth by convention and seems plausibly illustrated in contemporary set theory'.

This attitude to set theory would be opposed by Henryk Mehlberg ('The Present Situation in the Philosophy of Mathematics') who observes that 'abstract set-theory has been axiomatised in . . . many ways which are obviously incompatible with each other' (p. 79). There cannot be incompatible truths, even by convention, and since it is clear that Quine does not mean that one particular contemporary theory of sets is *the* true one, Mehlberg's view must be regarded as opposed to Quine's. Quine's view might be defended by someone who held that modern theories of sets involve 'legislative definition' (p. 50) so that the various theories don't mean the same thing by 'set'. But then the theories wouldn't be incompatible, so the disagreement with Mehlberg couldn't be resolved in this way.

In 'Epistemology and Logic', Abraham A. Fraenkel discusses the Continuum Hypothesis in the light of data (since extended) showing it to be independent. The suggestion is that this mathematical data alone cannot determine whether the hypothesis is true, but that the question of the truth of the hypothesis still arises. Hence Fraenkel would also seem to be opposed to the suggestion that this hypothesis of set theory is true (or false) by convention and thus opposed to the view of Quine's just mentioned, at least for this part of set theory.

I cannot resist taking space to register a disagreement of my own with one of Quine's points. He says :

> Suppose it claimed that . . . natives accept as true a certain heathen sentence of the form 'q ka bu q' the English translation of which has the form 'p and not *p*'. . . . just how good a translation is this, and what may the lexicographer's method have been? If any evidence can count against a lexicographer's adoption of 'and' and 'not' as translations of 'ka' and 'bu', certainly the native's acceptance of 'q ka bu q' counts overwhelmingly. (p. 41).

Now suppose we have translated something of the form 'q ka bu q' as 'We sometimes step into the same rivers and it is not the case that we sometimes step into the same rivers', and we have good evidence that the tribe has a mystical religion. Must we start looking for another translation just because the tribespeople all accept the sentence in question as true?

Jørgen Jørgensen, in 'Some Remarks Concerning Languages, Calculuses, and Logic', argues that the content of language is concepts and propositions and that logic is about the content of language rather than language alone.

Space does not permit me to go beyond the foregoing very sketchy indication of the contents of this book, but it should at least serve to show that the occasion of Carnap's seventieth birthday has inspired an impressive tribute from a number of famous philosophers and logicians.

JAMES CARGILE

Science, Technology, and Human Values. By A. Cornelius Benjamin. University of Missouri Press, Columbia, Mo., 1965. Pp. x+296. $7.00.

THIS book apparently is addressed to those who have had little training in science, technology, or philosophy. It may not be appropriate to say 'addressed to laymen' or 'addressed to non-scientists', for so broad is Professor Benjamin's definition of science that the class of non-scientists may provide too tiny an audience for any non-fictional book. Although the author twice warns us, 'When anything calls itself science, beware!' he bestows the title with princely largesse, never indicating that most of his references would omit from the category *science* not only such a discipline as pure mathematics, but jurisprudence, theology, advertising, and politics as well. Described are such scientific experiments as the passing of prohibition laws, marrying, divorcing, joining the common market, etc.; even the ascertaining in 1939 of what people would do if Germany were to invade Poland was in a sense an experiment. In such a context, it is almost reassuring that no emphasis is placed on reproducibility, recurrence, or even predictability as criteria for scientific experimentation.

The book is principally an exercise in pragmatic metaphysics, focusing on values and specifically the value of inquiry. Philosophically, science is one of several value pursuits, not different in justification from art, religion, morality, work or play; in its ascendancy today, science tends to force on all society its view of truth and learning as the highest values. The author succeeds best when treating interdisciplinary correlations, emphasising in novel ways the scientific aspects of a variety of endeavours not commonly thought of as scientific. He also presents unusual treatments of basic topics: on observation and getting the facts, on the language of science, on formal science, on metasciences, and on the relationship of philosophy to science. Valuable features are the use of tables and diagrams to make clear relationships which are of especial importance to the argument and a good bibliography which is presented as references. A chapter on the scientist's habit of thought is well summarised by its clever subtitles: Fervent Impartiality, Imaginative Objectivity, Tolerant Rigidity, and Critical Illogicality.

The writing, smooth and interesting, accomplishes the aim of avoiding technical language (and technical rigour?), but lapses too comfortably into vacuous generalisations: 'So far as science is concerned we shall be concerned only with such of these symbols as are capable of being interpreted by human beings.' 'Most people who have had sudden insights find it difficult to doubt them . . .' 'One frequently sees signs, " Practising Metaphysician ", on office doors . . .' and 'Most people do not realize how recently the word " science " was added to the English language.' 'Science' (rather than 'scientist') we are told was coined in 1840 but not included in the *Oxford English Dictionary* until 1867 ; the first date is at least 500 years too late for English, the second over forty too early.

The presentations of some aspects of the scientific enterprise are vague. In a section on the unsolved problem of insight, a discussion of flashes of genius is sandwiched between paragraphs on serendipity, thus glossing over any distinction between the mental climax which leads to the solution of a conscious problem and the unexpected discovery of something not sought for. In discussing flashes of insight, good use is made of quotations, a device which might have been used with profit at other places without compromising the freshness of the approach. The remarks on objectivity and the scientist are idealistic. Benjamin, unlike Conant or Snow, apparently feels that a scruple of honesty which hangs on the fear of being recognised as dishonest is yet a scruple.

Here and there straw men are erected. The current opinion of the common man that science is a remote and academic affair devoted to useless research is presented as somehow derived from the attitude of an earlier day when science was but the avocation of cultured gentlemen. Secrecy in science, specifically industrial and governmental science, Benjamin finds has resulted in the abandoning largely of a code requiring the immediate publishing of the results of research. He fails to note the flood of published results which today threatens to swamp publishers and frustrated readers alike.

The section on science and religion is unhackneyed and tolerant. However, science and religion are seen as so alike that, to paraphrase Wilde, both scientist and theologian are left with nothing to disbelieve in. One puzzling feature is the citing of an apparently sarcastic parody of the Apostles' Creed as a serious effort to establish the assumptions of science by the techniques of science.

In the sphere of ethics, a formal section developing the idea that science and scientists *qua* scientists are morally neutral is illustrated with all the common contraries, poisons *versus* penicillin, etc. Yet at other places science is seen with the promise ' to eliminate evil from the world ', or to transform nature to provide ' less sin and vice '; further, if an individual ' wishes to be more religious, the applied science of ministerial training and proselytising will be of value to him . . .'.

This is a provocative book, but one which is difficult to recommend either to student or dilettante in the philosophy of science.

C. W. KEENAN

My View of the World. By E. Schrödinger. Cambridge University Press, 1960. Pp. viii+110. 18s.

THIS book consists of two essays, of which the second supplements the first, written in 1925 and in 1960 respectively by the very distinguished theoretical physicist, the late Dr Schrödinger.

The book does not deal in any way with the problems of the interpretation of wave mechanics which the author did so much to discover. It is concerned mostly with the relation between mind and matter, in particular the existence or not of an external world and the extent to which mind can survive death. After rejecting as ' altogether too naïvely puerile ', the idea ' of a soul dwelling in the body as in a house quitting it at death and capable of existing without it ', Schrödinger is led to the Indian view that all consciousness is one, ' the admission that we are all really only various aspects of the One '.

I am unfortunately not sufficiently well read in philosophy to know how far the arguments presented here are new, but they are attractively set out. They bear the imprint of a first-class intelligence, which is not surprising considering who was the author.

Schrödinger, who is not afraid to write at one point that ' metaphysics . . . even mysticism . . . play an essential part in all that follows ', believes that the idea of the oneness of consciousness has profound ethical value.

I find it difficult to avoid being influenced in these matters by experience of physics. The ideas of continuity and discontinuity both seem to me so essential to an understanding of the physical world that I am instinctively repelled by an attempt to reduce everything to a hypothetical ' one '. I prefer the Persians to the Indians, if we must be Oriental, and the duality of Zoroaster to the unity of Brahma. In physics Occam's razor, as most experimentalists would agree, is at best a blunt instrument capable only of clearing away what is quite irrelevant. Once a rough outline is established further knowledge shows that the simplicity of the original discovery is only an approximation.

This is a thoroughly readable book. It has certainly provoked thought in one physicist ; I expect and hope it will do so in many more.

G. P. THOMSON

Contributions to Logic and Methodology in Honor of J. M. Bocheński. Edited by Anna-Teresa Tymieniecka in collaboration with Charles Parsons, North Holland Publishing Company, Amsterdam, 1965. Pp. 326. 65 Gldrs.

ONE tends to think of Fr Bocheński as primarily a logician, and this volume of essays, compiled on the occasion of his sixtieth birthday, reflects this central aspect of his work, and not his more recent concern with Soviet thought. His publications, a list of which prefaces the essays presented, testify to the existence of this new phase, as well as to the extraordinary industry and enterprise of their author.

As is most fitting in a volume dedicated to one who has loudly proclaimed the inadequacy of the old Prantl-based history of logic, there are several essays which are devoted to historical matters. Benson Mates discusses and clarifies the ingenious paradoxes evolved by Pseudo-Scotus around the notion of the soundness of *consequentiae*, and Ivo Thomas calls attention to three sixteenth-century writers who dealt in detail with the *pons asinorum*—a method, which later appears to have become unfashionable, for discovering suitable syllogistic middle terms. The lengthy essay on presocratic explanation contributed by Jerry Stannard illuminates the attitude of the presocratics to inference, evidence, classification, hypotheses, and generalisation. He conveys the impression (e.g. on pp. 108-109) that for him science involves the *explicit formulation* of an explanatory method; but is not the *use* of such a method sufficient for science, and does not its explicit formulation belong rather to the realm of meta-science? In a similar vein he refuses to allow that Heraclitus violated the Principle of Contradiction; this he does on the grounds that Aristotle had not yet formulated that principle. However, right and wrong *uses* of language can be said to exist at a stage when the grammar of the language has not yet been formulated; otherwise there would seem to be little point in starting up the study of a language's grammar. Correspondingly, there seems to be no reason to refuse the name ' science ' to the use of what later becomes codified in a statement of method, or to eschew the description of inferential absurdities as breaches of logic, notwithstanding the non-existence of logical theory. G. L. Kline suggests that the Russian logician N. A. Vasil'ev should be credited with the origination of the general idea of a three-valued logic. He recounts how, in four papers published between 1910 and 1913, Vasil'ev made the interesting suggestion that one might have a ' non-Aristotelian ' logic in the same way as one has non-Euclidean geometry. Here, however, to judge by Kline's account, the interest ends. Where not completely confused, or involved in logically irrelevant considerations, Vasil'ev seems to be grappling with points which an adequate modern logic, duly extended to the metalinguistic level, could cover without being compelled to resort to three-valued logic. He is considerably hampered by his restriction to the limited and inexact framework of old-fashioned pseudo-Aristotelian syllogistic. In any case, as Kline admits in various footnotes, several earlier thinkers, Scotus and Ockham, for instance, have already been credited with logics other than two-valued logic.

In an admirably crisp and clear essay, R. M. Chisholm gives his solution to the paradox raised by belief-statements when Leibniz's Principle of the Identity of Indiscernibles is brought to bear. R. M. Martin, on the topic of logic and ontology, argues with energy against the notion that logic is a branch of mathematics. However, he totally ignores the useful support of his contention which was made by Leśniewski. I found F. R. Barbo's contribution to the sense and reference discussion somewhat obscure in places, and suspect that its immense range—it covers the works of Frege, Russell, Quine, and others—is such that misunderstandings can all too easily creep in. ' Structure ' is an undefined keyword in A.-T. Tymieniecka's *Conjectural Inference and Phenomenological Analysis*; the reader is hence left to guess wildly what the meaning of sentences such as the following may be: ' relying upon the separate individual structures it [i.e. the programme of enquiry proposed in the essay] seeks the outline of the whole, of the universal order and of the architectonic project. It is no more a new structure we look for, but the set of conditions for all structures given ' (p. 251). However, the last section of this essay deals with conjectural thinking in the philosophy of Leibniz; the comparatively familiar subject-matter is here a great help to intelligibility, and we seem to have some quite interesting suggestions as to the way in which Leibniz worked.

The artificialisation, for philosophical and logical purposes, of natural language is, as J. F. Staal suggests in his *Reification, Quotation, and Nominalisation*, a phenomenon of great importance and interest. His method of investigating the tranformational rules observed in the artificialisation of Sanskrit could, I suspect, serve as a model for a most profitable piece of research into the similar type of artificialisation which occurs in medieval Latin, and to which current European philosophical language owes so much.

About one-quarter of the whole book is occupied by studies in Natural Deduction by P. Bernays, H. B. Curry, and H. Leblanc. Other contributors are S. McCall (on modal logic), A. Menne, H. S. Houthakker, N. M. Luyten, and E. W. Beth. Typographically, the volume is beautifully produced, and appears to contain only a few minor printing errors.

DESMOND PAUL HENRY

Le Concept d'Information dans la Science Contemporaine. Cahiers de Royaumunt. Edited by M. Louis Couffignal. Gauthier-Villars, Paris, 1965. Pp. 423. 35F.

As a contribution to information this book makes a bad start, because it contains no data on place or time. One can only guess that there was a symposium with invited guests in a place called Royaumont (town, village, chateau?) certainly before September 1964, because Norbert Wiener was still alive, more likely in 1962, because reports of conferences do not come out so fast. A few bits of information would have made all the difference! One knows a little about the participants, the mathematicians Couffignal, Mandelbrot, Wiener, for instance, or the historian Santillana, but it would have been interesting to know for example, what subject Professor Greniewski is teaching in the University of Warsaw, or Professor Goldmann in Paris?

For my part, I was most impressed by what the mathematician Benoit Mandelbrot said in the third of the talks: 'I believe that if the concept of information is conceived in precise mathematical terms, the main direct consequences of its theory can be considered as already established. The theory had successes, later it ran up against unsurmountable conceptual difficulties, now it has " stepped down " a little. But still it keeps suffering from a multitude of emotional implications arising from its name, and from the truly extraordinary publicity which it had at its beginnings. This is such a handicap that one must consider its scientific utility as defunct at least for the time being.' And late on: 'Having heard the previous speakers, and having taken part in many similar meetings I have the solid impression that such meetings are no longer useful. . . . Information in its well-defined meanings has been sufficiently exploited . . . What remains is so difficult that one can discuss it usefully only in private.'

Needless to say that this did not stop the symposium, and it would have been regrettable if it had, because at least some of the speakers made interesting contributions. The most informative was perhaps Helmar Frank's talk on Information and Pedagogy. It appears now as established that the mind can take in a maximum of about 16 bits per second and store them for a maximum of 10 seconds in a ' short term store ', from where information cannot pass at more than a rate of 0·5 to 0·7 bits per second into the permanent memory. This leads to certain clear conclusions such that the academic teacher must use very short sentences with preference, and if he insists on a machine-gun delivery, he must pad his information with at least 96 per cent redundancy. Evidently there is still much research to be done on the best type of padding.

Another interesting technical contribution was that by R. de Possell on the problem of character recognition, which has become so important for the input of computors. Andre Lwoff gave a simple and clear exposition of information in molecular biology, and Abraham Moles of the information theory of perception. I was less impressed by the philosophical discussions of Goldmann and Zeman. Norbert Wiener talked on one of his familiar subjects, Man and Machine, in the now familiar way. In the discussion he disclosed that he was now very interested in the general problem of the relation of structure and function. One may regret that this capital problem came to him so late in life; it is likely to remain an open field for some generations. If ever a general algorithm could be found to translate function into structure, this would be the end of invention, which the referee for one would very much regret.

The discussions which followed the talks and which are reproduced verbatim (probably with some French polish applied to the contributions of the foreigners), contain many interesting remarks by participants who did not give lectures, but who are known for their critical talent and ready wit, like D. M. MacKay and W. A. Rosenblith.

On the whole, if one compares this symposium with the first international symposium on Information Theory in London in 1950, one cannot avoid a sense of disappointment. But this need not always remain so. In some years perhaps we shall obtain some real insight into information processing by the nervous system, and the programme outlined by J. v. Neumann shortly before his death, that ' one ought to treat automata thermodynamically ' may be at least partly realised. But until then, as Mandelbrot said, these problems will remain too refractory for public meetings. In these matters, as in so many others, there is no substitute for hard brain work!

D. GABOR

Atlas Publ. & Distr. Co. Ltd., London - Stechert-Hafner, Inc., New York - Hermann & Cie., Paris - Nicola Zanichelli, Bologna - H. Bouvier u. Co., Bonn a/Rh. - Mezhdunarodnaja Kniga, Moscou - Gerold & Co., Wien - F. Machado & Cia, Porto - Libreria Editorial Augustinus, Madrid - Librairie Payot, S. A., Lausanne - The Maruzen Company, Tokyo.

"SCIENTIA"

INTERNATIONAL REVIEW OF SCIENTIFIC SYNTHESIS

A REVIEW COVERING ALL SCIENCES

(1966 - 60th year)

Scientific Committee:
G. ABETTI - G. B. BONINO - P. CALDIROLA - L. CALIFANO - R. CANESTRARI - G. COLONNETTI - B. DE FINETTI - M. GORTANI - G. LEVI DELLA VIDA - G. MONTALENTI - E. PADOA - A. SESTINI - V. SOMENZI - S. TONZIG.

"SCIENTIA" is the only review of its type which has a world wide circulation : deals with the most recent and fundamental problems of all branches of science : can boast among its contributors of the most illustrious men of science in the whole world : publishes the articles in the native language of their Authors (English, French, Italian, German, Spanish). Each issue includes a **Supplement** *containing the complete French translation of the articles which in the text are published in language other than French.*

"SCIENTIA" has therefore a very strong appeal to the scientific-minded reader all over the world.

Full details and a free back copy will be sent by applying to
"SCIENTIA" - Via Donizetti, 55 - Milano
sending 5s. or equivalent amount in other currency, preferably in air-mail postage stamps of your country, merely to cover packing and postage. For a number of the current year, please send 20s. or equivalent amount in other currency.

ANNUAL SUBSCRIPTION: £11.10s. 0d.

VOLUME VIII JUNE 1966 No. 1

RATIO

edited in co-operation with
H. B. Acton (London), P. Bernays (Zürich), A. Church (Princeton), H. Cherniss (Princeton), G. Henry-Hermann (Bremen), W. C. Kneale (Oxford), E. Nagel (Columbia), K. R. Popper (London) and J. W. N. Watkins (London)
by
Stephan Körner (Bristol)

W. STEGMÜLLER	Explanation, prediction, scientific systematization and non explanatory information
W. BEISSWANGER	Hermann Weyl and Mathematical texts
N. RESCHER and J. ROBINSON	Temporarily Conditioned Descriptions
J. O. WISDOM	Testing a Psycho-Analytic Interpretation
H. KHATCHADOURIAN	About Imaginary Objects
J. KNOX	On Mr. Nelson's Rejection of Sense-Data
A. FLEW	Did Hume Distinguish Pure from Applied Geometry?
Reviews:	
R. HARRÉ	*Matter and Method* (G. Buchdahl)
M. WARNOCK	*The Philosophy of Sartre* (D. Pole)

Price 15s. net **Annual Subscription: 27s. 6d. net**

BASIL BLACKWELL · OXFORD

JOURNAL OF THE HISTORY OF IDEAS

A Quarterly devoted to Intellectual History

VOLUME XXVII JAN.-MARCH 1966 NUMBER 1

TABLE OF CONTENTS

ARTICLES

Philosophy and Scientific Methods	*Richard McKeon*
Buridan and Seneca	*James J. Walsh*
Locke on Personal Identity	*Henry E. Allison*
Problem of Gulliver's Fourth Voyage	*John J. McManmon*
Idea of Westernizing Russia, 1830-1850	*Raymond T. McNally*
Malthus and the Idea of Progress	*Samuel M. Levin*
The Emergence of Progressive History	*Charles Crowe*

NOTES

Heraclitus and Fire	*Constance I. Smith*
T. Hariot: Early Atomism in England	*Robert Kargon*
J. G. Hamann: Metacritic of Kant	*W. M. Alexander*

REVIEW

Nakamura's *Ways of Thinking of Eastern Peoples*
reviewed by *A. T. Embree, C. M. Schirokauer, A. E. Tiedemann*

Subscription Rates (outside U.S.A. post paid): $6.50 for one year;
$1.75 (plus postage) for single issue

Payable to

JOURNAL OF THE HISTORY OF IDEAS, INC.
The City College, New York, N.Y. 10031

FOUNDATIONS OF LANGUAGE

International Journal of Language and Philosophy

Editorial Board - Morris Halle (MIT), Peter Hartmann (Münster/W), K. Kunjunni Raja (Madras), Benson Mates (Univ. of California), J. F. Staal (Amsterdam), Pieter A. Verburg (Groningen), John W. M. Verhaar (Manila)

VOLUME II No. 1 FEBRUARY 1966

Articles R. H. Robins: The Development of the Word Class-System of the European Grammatical Tradition — Charles Hartshorne: Some Reflections on Metaphysics and Language — D. Terence Langendoen: A Note on the Linguistic Theory of M. Terentius Varro — Charles E. Caton: On the General Structure of the Epistemic Qualifications of Things said in English — J. F. Staal: Analyticity

Reviews *Publications Received for Review*

FOUNDATIONS OF LANGUAGE is published quarterly

Annual subscription price *f* 42.—($11.75; 84s.)

D. REIDEL PUBLISHING COMPANY/DORDRECHT-HOLLAND